

# ICAETA -2021

International Conference on Advanced Engineering,  
Technology and Applications



# PROCEEDINGS BOOK 2021

## Host

Department of Computer Engineering,  
Istanbul Aydin University,  
Istanbul, Turkey.  
<https://izu.edu.tr>



**ISTANBUL AYDIN  
UNIVERSITY**

## Sponsors

AIPLUS  
308 Daisyfield Centre, Appleby Street  
Blackburn Lancashire BB1 3BL,  
United Kingdom  
<https://www.aiplustech.org/>





# **BOOK OF PROCEEDINGS**

**International Conference on Advanced Engineering, Technology  
and Applications (ICAETA -2021)**

**July 09-11, 2021,**

**Istanbul, Turkey**

## **Editorial Board**

Dr. Akhtar Jamil

Dr. Alaa Ali Hameed

**ISBN: 2752-8340**

Istanbul Aydın University

Beşyol, İnönü Cd. No:38, 34295 Küçükçekmece/Istanbul

Tel: 444 1 428

[www.aydin.edu.tr](http://www.aydin.edu.tr)



## Copyright © 2021

All Rights Reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, scanning or otherwise, except under the terms of the Copyright.

The individual contributions in this publication and any liabilities arising from them remain the responsibility of the authors. The publisher is not responsible for possible damages, which could be a result of content derived from this publication. Moreover, the proceedings book is published by AIPLUS in Blackburn, United Kingdom.

[info@aiplustech.org](mailto:info@aiplustech.org)

<https://icaeta.aiplustech.org>

## Keynote Speakers

### **Prof. Dr. Levent Sevgi**

IEEE Fellow, IEEE AP-S Distinguished Lecturer

Istanbul OKAN University,

Istanbul, Turkey

### **Prof. Dr. Ali Okatan**

Istanbul Aydin University.

Istanbul, Turkey

## Committees and Scientific Boards

### **General Chairs**

Prof. Dr. Naim Mahmood Musleh, Istanbul Aydin University, Istanbul, Turkey.

### **Technical Program Chairs**

Dr. Akhtar Jamil, Istanbul Sabahattin Zaim University, Turkey.

Dr. Alaa Ali Hameed, Istanbul Sabahattin Zaim University, Turkey.

Dr. Hasan Alpay Heperkan, Istanbul Aydin University, Istanbul, Turkey.

Dr. Bahar Ferah, Istanbul Sabahattin Zaim University, Turkey.

Dr. Omar Algburi, Istanbul Sabahattin Zaim University, Turkey.

### **Publication Committee**

Dr. Imran Ahmed Siddiqi, Department of Computer Science, Bahria University, Pakistan.

Maryam Torabi, Alzahra Technical and Vocational College, Iran

Dourna Kiavar, Sama University of Tabriz, Iran

Mustafa Takaoglu, Istanbul Aydin University, Turkey.



### **Registration Committee**

Dr. Amani Yahyaoui, Istanbul Sabahattin Zaim University, Turkey.

Alireza Sakha, Islamic Azad University, Iran

Ali Khiabani, Interdisciplinary Design Universe Office, Iran

Ayşe Gul Gemci, Istanbul Technical University, Turkey

### **Publicity Committee**

Dr. Adem Ozyavas, Istanbul Aydin University, Turkey.

Sama Khattab, Istanbul Aydin University, Turkey.

### **Program Committee**

Dr. Ameer Al-Nemrat, School Of Architecture Computing And Engineering, University Of East London

Dr. Chawki Djeddi, Laboratoire D'informatique De Traitement De L'information Et Des Systemes (Litis), University Of Rouen, France.

Dr. Can Balkaya, Department Of Civil Engineering, Istanbul Aydin University, Istanbul, Turkey.

Dr. Muhammad Abdul Basit, Montana Technological University, Butte Montana, Usa.

Dr. Naghmeh Moradpoor, School Of Computing, Edinburgh Napier University, United Kingdom

Dr. Kamran Dehghan, Department Of Architecture, Islamic Azad University, Iran

Dr. Selda Nazari, Department Of Architecture, Islamic Azad University, Iran

Dr. Syed Attique Shah , Balochistan University Of Information Technology Engineering And Management Sciences, Pakistan

Dr. Müberra Eser Aydemir, Department Of Civil Engineering, Istanbul Aydin University, Istanbul, Turkey.

Dr. Muhammad Fahim , Institute Of Information Security And Cyberphysical Systems, Innopolis University, Russia.

Dr. Shahab Adam Navasi, Department Of Architecture, Islamic Azad University, Iran

Dr. Muhammad Ilyas, Department Of Electrical And Electronics Engineering, Altinbas University, Turkey.

Dr. Rawad Hammad, School Of Architecture, Computing And Engineering, University Of East London, United Kingdom

Dr. Prateek Agrawal, Department Of Computer Science, University Of Klagenfurt, Austria.

Dr. Atoosa Modiri, Department Of Architecture, Islamic Azad University, Iran

Dr. Hasan Volkan Oral, Department Of Civil Engineering, Istanbul Aydin University, Istanbul, Turkey.

Dr. Jayapandian N, Department Of Computer Science And Engineering, Christ University, Bangalore, India.

- Dr. Kaveh Dehghanian, Department Of Civil Engineering, Istanbul Aydin University, Istanbul, Turkey.
- Dr. Nidaa Flaih Hassan, Department Of Computer Science, University Of Technology, Iraq
- Dr. Amani Yahyaoui, Department Of Computer Engineering, Istanbul Sabahattin Zaim University, Turkey.
- Dr. Firas Ajlouni, Department Of Computer Science, Lancashire College, United Kingdom.
- Dr. Bita Bagheri, Department Of Architecture, Islamic Azad University, Iran
- Dr. Zeynep Kerem Öztürk, Department Of Interior Architecture And Environmental Design, Istanbul Sabahattin Zaim University, Istanbul, Turkey
- Dr. Aliyu Musa, Predictive Society And Data Analytics Lab, Faculty Of Information Technology And Communication Sciences, Tampere University, Tampere, Finland.
- Dr. S Amutha, Department Of Computer Science And Engineering, Sàveetha Engineering College (Autonomous), Affiliated To Anna University, Chennai, India.
- Dr. Sepanta Naimi, Department Of Civil Engineering, Istanbul Aydin University, Istanbul, Turkey.
- Dr. Mehmet Fatih Altan, Department Of Civil Engineering, Istanbul Aydin University, Istanbul, Turkey.
- Dr. Mohammad Golmohammadi, Department Of Architecture, Islamic Azad University, Iran
- Dr. Lisa Oliver , Department Of Computer Science, Lancashire College, United Kingdom.
- Dr. Ahmet Gürhanli, Department Of Computer Engineering, Istanbul Aydin University, Istanbul, Turkey.
- Dr. Daniel White, Department Of Mathematics And Computing , Lancashire College, United Kingdom.
- Dr. Yasmin Doozdoozani, Department Of Architecture, Islamic Azad University, Iran
- Dr. Javad Eiraji, Faculty Of Architecture And Design, Eskisehir Technical University, Turkey
- Dr. Kiarash Eftekhari, Department Of Architecture, Islamic Azad University, Iran
- Dr. Mina Najafi, Editorial Assistant, Emerald Publishing Ltd, United Kingdom.
- Dr. Nahid Khahnamouei, Department Of Architecture, University Of Nabi Akram, Iran
- Dr. Murtaza Farsadi, Department Of Electric-Electronic Engineering, Istanbul Aydin University, Istanbul, Turkey.
- Dr. Sadeq Alhamouz, Department Of Computer Sciences, Wise University, Jordan.
- Dr. Necip Gökhan Kasapoğlu, Department Of Electric-Electronic Engineering, Istanbul Aydin University, Istanbul, Turkey.
- Dr. Vira V. Shendryk, Department Of Computer Sciences, Sumy State University, Ukraine.
- Dr. Elnaz Pashaei, Department Of Software Engineering, Istanbul Aydin University, Istanbul, Turkey.
- Dr. Shareeful Islam, School Of Architecture Computing And Engineering, University Of East London, London, United Kingdom
- Dr. Navid Khaleghimoghaddam, Department Of Engineering And Architecture, Konya Food And Agriculture University, Turkey.
- Dr. Nima Mirzaei, Department Of Industrial Engineering, Istanbul Aydin University, Istanbul, Turkey.

Dr. Imene Yahyaoui, Universidad Rey Juan Carlos, Applied Mathematics Materials Science And Engineering And Electronic Technology, Spain.

Dr. Ilham Huseyinov, Department Of Software Engineering, Istanbul Aydin University, Istanbul, Turkey.

Dr. Azeem Hafeez, Department Of Electrical And Computer Engineering, University Of Michigan, Usa.

Dr. Abdulkader Alwer, Department Of Electric-Electronic Engineering, Istanbul Aydin University, Istanbul, Turkey.

Dr. Zeynep Orman, Department Of Computer Engineering, Istanbul University, Turkey.

Dr. Xiaodong Liu, School Of Computing, Edinburgh Napier University, United Kingdom

Dr. Saed Moghimi, Department Of Civil Engineering, Istanbul Aydin University, Istanbul, Turkey.

Dr. Alaa Sheta, Department Of Computer Sciences, Southern Connecticut State University, Usa.

Dr. Mehmet Güneş Gençyilmaz, Department Of Industrial Engineering, Istanbul Aydin University, Istanbul, Turkey.

Dr. Raheleh Mirzaei, Department Of Industrial Engineering, Istanbul Aydin University, Istanbul, Turkey.

Dr. Vijayakumar Varadarajan, School Of Computer Science And Engineering, The University Of New South Wales, Sydney, Australia.

Dr. Luca Romeo, Istituto Italiano Di Tecnologia, Italy.

Dr. Numan Khurshid, Seecs, National University Of Science And Technology, Pakistan

Dr. Elif Merve Kahraman, Department Of Food Engineering, Istanbul Aydin University, Istanbul, Turkey.

Dr. Sibel Kahraman, Department Of Food Engineering, Istanbul Aydin University, Istanbul, Turkey.

Dr. Mustafa Nafiz Duru, Department Of Industrial Engineering, Istanbul Aydin University, Istanbul, Turkey.

Dr. Saeid Homayouni, Centre For Water, Earth, And Environment, Inrs-Quebec, Canada

Dr. Imran Ahmed Siddiqi, Department Of Computer Science, Bahria University, Pakistan.

Dr. Zeynep Dilek Heperkan, Department Of Food Engineering, Istanbul Aydin University, Istanbul, Turkey.

Dr. T. Lalitha, Computer Science & It, Jain Deemed-To-Be University, Bengaluru, Karnataka, India

Dr. Aysa Jafari Farmand, Istanbul Technical University, Turkey.

Dr. Mehdi Zahed, School Of Applied Sciences And Technology, Nait, Canada.

Dr. Mehmet Güneş Gençyilmaz, Department Of Food Engineering, Istanbul Aydin University, Istanbul, Turkey.

Dr. Gülay Baysal, Department Of Food Engineering, Istanbul Aydin University, Istanbul, Turkey.

Dr. Fahimeh Jafari, School Of Architecture, Computing And Engineering, University Of East London, United Kingdom

Dr. Mustansar Ali Ghazanfar, Department Of Computer Science And Digital Technologies, University Of East London, United Kingdom

Dr. Sibel Senan, Department Of Computer Engineering, Istanbul University, Turkey.





Dr. Mehmet Emin Tacer, Department Of Electric-Electronic Engineering, Istanbul Aydin University, Istanbul, Turkey.

Dr. Mohammed Alkrunz, Department Of Electrical And Electronics Engineering, Istanbul Aydin University, Istanbul, Turkey.

Dr. Mohammed Vadi, Department Of Electrical And Electronics Engineering, Istanbul Sabahattin Zaim University, Istanbul, Turkey.

## TABLE OF CONTENTS

S.No	Autors/Titles	Page
1	<b>Mesut Toğaçar and Burhan Ergen</b> Recognition Human Activities by Convolutional Based Bidirectional LSTM Model Using ECG Signal Data	1
2	<b>Adel Ridha Othman</b> Preventing Oscillation of Supply Voltage Due to Resonance Harmonic Frequency of a Power Factor Corrected System	6
3	<b>Parastoo Pourvahidi</b> Arranging Spaces for Next Pandemic by inspiring from Qajar period	11
4	<b>Adeola Opesade</b> Comparison of Sentiment-Lexicon-based and Topic-based Sentiment Analysis Approaches on E-Commerce Datasets	18
5	<b>Abdulazeez Sulaiman, Fatih Özyurt and Shivan Mohammed</b> Autism detection from facial images using deep learning methods	24
6	<b>Indronil Bhattacharjee, Al Mahmud and Tareq Mahmud</b> Diabetic Retinopathy Classification from Retinal Images using Machine Learning Approaches	29
7	<b>Melda Ozdinc Carpinlioglu</b> Methodology for Syngas Energy Assessment	34
8	<b>Mohamed K. Hussein, Mahmoud I. Khalil and Bassem A. Abdullah</b> 3D Object Detection using Mobile Stereo R-CNN on Nvidia Jetson TX2	39
9	<b>Aida Ghiaseddin and Javad Eiraji</b> Investigation of Algorithmic Architecture Design Method by Using Digital Technology to Increase Flexibility in Design Process	44

<b>10</b>	<b>Huwaida Elgweri, Amal Hamed and Mohamed Mansor</b> Applying of the Finite Difference Time Domain Method for the Calculation of Lower Angular Excited States of Symmetrical Two Dimensional System	49
<b>11</b>	<b>Zeba Mahmood</b> Implementation Framework for a Blockchain-Based Reputation and Trust System	55
<b>12</b>	<b>Edmond Manahasa and Rudina Kazazi</b> An Observation on Residential Complexes as a New Housing Typology in Post-Socialist Tirana	62
<b>13</b>	<b>Abbas H. Hammood and Başar Özkan</b> Active Control Of In-Wheel Motor Electric Vehicle Suspension Using The Half Car Model	68
<b>14</b>	<b>Lamiya Salman</b> Evolutionary Deep automatic CAD system for Early detection of Diabetic Retinopathy and its severity classification.	75
<b>15</b>	<b>Oguzhan Erbas, Ahmet Akbulut and Fadime Menekşe İkbal</b> Studies on Increasing Energy Efficiency By Modernization Of A Single-Stage Type Turbo-Compressor For Ammonia Combustion Air Process	81
<b>16</b>	<b>Wafa Guedri, Mounir Jaouadi and Slah Msahli</b> Prediction of Cover Bunch Quality After Harvesting Period of Tunisian Date Palm	87
<b>17</b>	<b>Anna Yunitsyna and Mirela Hasanbashaj</b> Analysis of the Waterfront Transformation of the ‘Plazh’ Area of the City of Durres, Albania	91
<b>18</b>	<b>Şafak Kılıç, İman Askerzade and Yılmaz Kaya</b> Deep Learning Using MobileNet for Personal Recognizing	97

19	<b>Kübra Kiliçarslan, Emirhan Sağ and Abdurrahman Günday</b> Effects of Photon-Shot and Excess Noises on Detectable Minimum Rotation Rate in I-FOG Design for Autonomous Vehicles	103
20	<b>Spencer Li</b> How Does the after-COVID-19 “ABCDEF” effects model affect the development of Internet of Things and its Applications to improve Customer Experiences?	108
21	<b>Jara Muñoz-Hernández</b> Drawing as a Scientific Method. The School of Agricultural Engineers in Madrid: a case study	114
22	<b>Juan Pablo Vasco-Gallo, J. Isaac Pemberthy-R. and Eduard A. Gañan-Cardenas</b> Urban distribution network proposal: A case study for the 14th district of the city of Medellín.	120
23	<b>Purushoth Anantharasa and Ragu Sivaraman</b> Close Price Prediction of Day Stock Markets with Machine Learning and NLP models	125
24	<b>Kaveh Dehghanian and Mohammad Haroon Saeedi</b> Investigation of Permeability Coefficient in Layered Soils	131
25	<b>Munya Alkhalifa and Kasım Özacar</b> An AI-based Embodied Digital Human Assistant for Information in University	137
26	<b>Hanan Othman, Yakoub Bazi and Mohamad Alrahhal</b> Visual Question Answering for Medical Image Analysis based on Transformers	143
27	<b>Tuba Elmas Alkhan, Alaa Ali Hameed and Akhtar Jamil</b> Deep Learning for Face Detection and Recognition	148

28	<b>Muhammad Farooq, Ali Javed, Khalid Malik and Anas Raza</b> A Lightweight and Interpretable Deepfakes Detection Framework	154
29	<b>Edmond Manahasa and Arila Rasha</b> Searching for Aesthetical Values in an Upgraded Informal Neighborhood in Tirana	160
30	<b>Oğuzhan Erbas, F.Menekşe İkbāl and Ahmet Akbulut</b> Effect of Oxidation Reactor Structure on Operating Parameters and System Performance in a Nitric Acid Production Plant	166
31	<b>Önsen Toygar, Mehtap Köse Ulukök and Emre Özbilge</b> Plant Disease Identification Through Deep Learning	171
32	<b>Abdulkadir Şahiner, Kaan Kemal Polat and Hayati Ünsal Özer</b> Vaccines Perspective in the COVID-19 Era: Analysis of Twitter Data	176
33	<b>Mohammad Ikhsan Zakaria, Akhtar Jamil, Alaa Ali Hameed</b> Deep Learning-based Healthcare Data Analysis System	181
34	<b>Sami Benni and Nildem Tayşi</b> Optimum Usage of Mixed Damping Systems (Rubber Concrete or X Diagonal Dampers) on Multistory Building	186
35	<b>Murat Mustafa Savrun</b> Bidirectional DC-DC Converter Based on Quasi Z-Source Converter with Coupled Inductors	193
36	<b>Fabio Naselli and Enkela Krosi</b> Influences of urban fabrics onto microclimate assessment within the city of Tirana	198
37	<b>Sarosh Ahmad, Sheza Yasin, Sajal Naz, Amina Batool and Ali Suqrat</b> IoT Based Water Management and Monitoring System for Multi-Resources	202

38	<b>Sarosh Ahmad, Arslan Dawood Butt, Usama Umar, Sajal Naz, Sheza Yasin and Amina Batool</b> Development of a High Precision Temperature Monitoring System for Industrial Cold Storage	208
39	<b>Mustafa Nurmammed, Ozan Akdağ and Teoman Karadağ</b> Modeling and Load Flow Analysis of Electric Vehicle Charging Stations in Power Distribution Systems	213
40	<b>Mohammed Majid Abdulrazzaq, Mustafa Mohammed Alhassow and Abdullah Ahmed Al-Dulaimi</b> Obstacle Avoiding Capabilities for The Drone by Area Segmentation and Artificial Neural Network	218
41	<b>Muhammad Fawad Khan and Muhammad Azam</b> Joint User Selection and Base Station Assignment Strategy in Smart Grid	223
42	<b>Mert Safak Tunalioglu</b> Comparison Of Flowdrill And Conventional Drilling Methods In Thin-Walled Materials	228
43	<b>Shahram Taheri, Zahra Golrizkhatami and Önsen Toygar</b> Classification of Animal Faces Using A Novel DAG-CNN Architecture	233
44	<b>Bugay Sarikaya and Duygu Dede Sener</b> Movie Success Prediction with Statistical Analysis Techniques and Machine Learning Methods	239
45	<b>Abdullah Ahmed Al-Dulaimi, Mohammed Majid Abdulrazzaq, Mustafa Mohammed Alhassow and Noor Qasim Atiyah Al Saedi</b> Digital Controllers Performance Analysis for a Robot Arm	244
46	<b>Md Riyad Hossain, Douglas Timmer and Hiram Moya</b> Machine Learning optimization with Hyper-parameter tuning approach	250

<b>47</b>	<b>Pınar Terzioglu and Eslem Kavas</b> Polyvinyl Alcohol/Cassava Starch/Nano-CaCO <sub>3</sub> based Nano-Biocomposite Films:Mechanical and Optical Properties	256
<b>48</b>	<b>Nour Ammar, Ali Okatan and Naim Ajlouni</b> Movie Reviews Text Sentiment Analysis Based on Hybrid LSTM And Glove	260
<b>49</b>	<b>Omar Algburi and Bahar Ferah</b> Simulation Comparative Study to Highlight the Relation Between Building Form and Energy Consumption	265
<b>50</b>	<b>Osen Fili Nami, Abdul Halim and Dewi H. Budiarti</b> Comparison of PID and LQR Controller of Autonomous Underwater Vehicle for Depth Control	270
<b>51</b>	<b>Büşra Nur Darendeli and Alper Yılmaz</b> Convolutional Neural Network Approach to Distinguish and Characterize Tumor Samples Using Gene Expression Data	276

# Recognition Human Activities by Convolutional Based Bidirectional LSTM Model Using ECG Signal Data

Mesut Toğaçar  
Computer Technologies Department  
Firat University  
Elazığ, Turkey  
mtogacar@firat.edu.tr

Burhan Ergen  
Computer Engineering Department  
Firat University  
Elazığ, Turkey  
bergen@firat.edu.tr

**Abstract**—Recognition of human activities is a challenging process and technology-based systems are used for the successful realization of the process. Recently, artificial intelligence-based technologies have started to be used widely. The hybrid approach designed for this study consists of convolutional-based bidirectional long short-term memory (C-BiLSTM). In this study, 12 types of human activities were identified using C-BiLSTM using ECG signal data. As a result of the analysis, an overall accuracy success rate of 98.96% was achieved. The result obtained in the experimental analysis has been promising in identifying the types of human activity.

**Keywords**— *bidirectional LSTM, deep learning, ECG measurement, physical activities, human behavior*

## I. Introduction

Today, recognizing human activities has become easier thanks to various technological approaches (wearable sensor, video, accelerometer, signal data, etc.) [1]. Information on the human activity actions is used in areas such as following the elderly, performing health actions correctly, criminal tracking systems, etc. [1], [2]. Technological infrastructure systems are supported by software containing innovative aspects [3]. Artificial intelligence technologies form part of the innovative aspects. Artificial intelligence-based systems offer a service where people can perform their transactions more easily today [4], [5].

Many studies have been conducted recognition artificial intelligence-based human activities. Shaohua Wan et al. [6] conducted the recognition of human activities using machine learning and deep learning methods. They used convolutional neural network (CNN), support vector machines (SVM), short long-term memory (LSTM), multilayer perceptron (MLP) approaches in their study. They achieved the best accuracy success rate of 92.71% with the CNN model. Junfang Gong et al. [7] used social media data to recognition human daily activities. They achieved an overall accuracy rate of 89.35% with the LSTM model they designed in their study. Nozha Jlidi et al. [8] used the transfer learning-based PoseNet model to recognition human activities. They successfully achieved accuracy by emphasizing the body joints. Emilio Sansano et al. [9] used gated recurrent unit networks (GRU), deep belief networks (DBN) and LSTM approaches to recognize human activities. They achieved over 90% overall accuracy success in all of the approaches they used. Sakorn Mekruksavanich et al. [10] proposed a biometric user identification-based approach to defining human activities for the health status monitoring of elderly people. They obtained data using a triaxle gyroscope and triaxle accelerometer. They used the CNN model and LSTM model in the analysis of the data. The classification success of the models was 91.77% and 92.43%, respectively. Negar Golestani et al. [11] presented a wireless approach based on magnetic induction to recognize human activities. They categorized their activity activities by integrating the magnetic induction system with machine

learning techniques and deep learning approaches. They achieved the best performance with the LSTM model. Accuracy success for the two datasets they used was 87% and 98.9%, respectively. S. Tsokov et al. [12] used accelerometer sensors to classify human activity types. They designed a 1D-CNN model to describe the data they obtained according to their types. The overall accuracy success they got from the CNN model they designed was 98.86%.

The goals of this study are; it is to successfully recognize human activities with the proposed approach. The proposed approach is hybridized CNN & Bidirectional LSTM models designed using python libraries. The summary of this article about the sections is as follows; in the second section, information about the dataset is given. The third section contains information about deep learning approaches and the proposed approach. The fourth section contains the results of the experimental analysis. The last section consists of information about Discussion and Conclusion.

## II. Dataset

The dataset consists of ECG data that includes various physical activities created with 10 volunteer participants. Using wearable sensors to measure movement data and vital signs, other feature data were also created to detect activities. Measurements were made with an average speed of 50Hz during the creation of ECG data. The activity types created are 13 in total [13], [14]. The repetition/duration equivalents of the activity types and types that make up the dataset are given in Table 1.

Table 1. Repetition/duration equivalents of the activity types and types that make up the dataset

Activity Type	Times (x) / minute (m)	Activity Type	Times (x) / minute (m)
Nothing	-	Frontal elevation of arms	20x
Standing still	1m	Knees bending	20x
Sitting and relaxing	1m	Cycling	1m
Lying down	1m	Jogging	1m
Walking	1m	Running	1m
Climbing stairs	1m	Jump front & back	20x
Waist bends forward	20x		

The dataset contains 14 feature columns for each data operation. In addition, 20% of the dataset was separated as test data in experimental analysis, and 80% was allocated as training data.



### III. Model Approaches

This section contains brief information about the models used in the proposed approach.

#### A. Bi-directional Long Short-Term Memory

The Bi-LSTM model is the combined state of bidirectional recurrent networks (Bi-RNN). The most important feature that distinguishes LSTM from RNN consists of gates used as memory units. Input data are processed through input gates and transferred to output gates. The difference between the Bi-LSTM model from the LSTM model is that it can return from the previous context and receive data. That is, in Bi-LSTM models, previous layer information and next layer information are kept in memory gates [15]. In Bi-LSTM models, the number of hidden units and their functioning are calculated according to Eq. (1) and Eq. (2). In these equations,  $L$  and  $H$  inputs are used for the number of hidden units. Here;  $t$  represents the time value,  $x^t$  is the sequence input,  $\theta_h$  the activation function of the hidden unit,  $w$  the weight values of the hidden unit, and the variable  $b^t$  represents the activation function of the  $h$  unit at time  $t$  [15], [16].

$$a_h^t = \sum_{l=1}^L x_l^t w_{lh} + \sum_{h'=1, t>0}^H b_{h'}^{t-1} w_{h'h} \quad (1)$$

$$b_h^t = \theta_h(a_h^t) \quad (2)$$

#### B. Convolution Neural Networks

CNN is an artificial intelligence-based model used in classification, recognition, segmentation, etc. processes by processing input data [17]. These networks generally consist of convolutional layers, pooling layers, fully connected/dense layers [18]. Apart from that, it can contain different layers in line with the target of the model. CNN models contain hidden layers in their architectural structures. Convolutional layers process the input data by filters to extract activation features [19]. The mathematical formula in Eq. (3) is used to extract the activation maps [20]. In this equation; the variable  $F$  represents the layer of the activation map. The variable  $n$  represents the number of features in the layer. Variables  $i, j$ , and  $k$  provide the position information of the input data. Matrix values are represented by variable  $M$ .

$$F_{i,l}^j = \sum_{k=1}^n M_{i,l}^{j,k} x_{k-1}^k \quad (3)$$

The pooling layer is often used after convolutional layers and helps to train the CNN model more easily by reducing the dimensions of the activation maps [21]. The mathematical formula in Eq. (4) is used to calculate the maximum pooling region. Here, the result of pooling is represented by a variable  $y_{kij}$ . The feature map is represented by  $k$ . The pooling region is represented by  $R_{i,j}$ , while the variable  $x_{kpq}$  is an element in this region. Also,  $(p, q)$  represents the position in the pooling region [20].

$$y_{kij} = \max_{(p,q) \in R_{i,j}} x_{kpq} \quad (4)$$

Fully connected layers collect the output values obtained from all layers in a single layer and help calculate probabilistic values for activation functions used in the classification process such as Softmax. Fully connected layers are often used in the last layers of CNN models [21], [22]. The Softmax function is generally used in the last layer of CNN models and is preferred due to its multi-classification feature. It extracts

the probability values as many as the type number of the features from the fully connected layer and tags the input to the dominant probabilistic type [23].

Also, the ReLU activation function is generally preferred between layers in CNN models. ReLU is an activation function that allows input values to be linearized and can keep negative values at zero [24], [25]. CNN can be used in batch-normalization and dropout functions to prevent models from being over / underfitting [26].

#### C. Proposed Approach

The proposed approach is the result of combining the CNN-based Bi-LSTM model, which aims to describe human physical activities. Using the activity action features and ECG signal data obtained from C-BiLSTM model sensors, it aimed to successfully classify 13 activity types. The C-BiLSTM model is completely designed in Python programming language, and "Tensorflow, Keras, Pandas, Numpy, etc." coding was carried out using libraries. Jupyter Notebook interface program was used in the compilation of the model. The general design of the model is given in Table 2. The proposed approach consists of layers of the CNN model and layers of the LSTM model. Since the two models are designed with open source codes, data transitions and parameter values between models in python software must be compatible. Therefore, in the proposed approach, the output values from the last layer of the CNN model were provided to be equal to the input values of the BiLSTM model. In other words, while providing the transition, the tensor numbers, parameter values, and input size are hybridized to be the same. The normalization between the feature values obtained from the convolutional layers was carried out with the batch - normalization layer. Also, using the dropout layer in the last layers of the proposed approach, inefficient features were prevented from being trained by the model. Thus, the model's training speed and time savings were achieved.

Table II. The general design of the proposed model

Layer	Value / Output Shape
Convolutional	(None, 400, 64)
Batch Normalization & ReLU	(None, 400, 64)
Convolutional	(None, 400, 128)
Batch Normalization & ReLU	(None, 400, 128)
Convolutional	(None, 400, 256)
Batch Normalization & ReLU	(None, 400, 256)
Bidirectional LSTM	(None, 400, 256)
Bidirectional LSTM	(None, 400, 128)
Maximum Pooling	(None, 128)
Dense	(None, 1024)
Dropout	0.25 / (None, 1024)
Dense	(None, 512)
Dropout	0.25 / (None, 512)
Dense	(None, 256)
Dropout	0.25 / (None, 256)
Dense	(None, 13)

In addition, in the measurements of the proposed model, the optimization method Rmsprop was preferred and the loss function was chosen as categorical cross-entropy.

#### IV. Experimental Analysis

Experimental analyzes were compiled on the Google COLAB server. Trainings were performed using the GPU and the preferred epoch value for training was 100. Other preferred parameters in model compilation; optimizer method rmsprop, loss function "sparse categorical cross-entropy" and metric function "sparse categorical accuracy" were selected. The "early stopping" parameter was used to prevent overfitting in the model. Confusion Matrix was used to compare analysis results. Eq. (5-8) were used to calculate matrix metrics. The variables used in these equations; (TP): true positive, (TN): true negative, (FP): false positive, and (FN): false negative are defined [27]–[29].

$$\text{Precision (Pre)} = \frac{TP}{TP+FP} \quad (5)$$

$$\text{Recall (Rec)} = \frac{TP}{TP+FN} \quad (6)$$

$$\text{F-scor (F-scr)} = \frac{2xTP}{2xTP+FP+FN} \quad (7)$$

$$\text{Accuracy (Acc)} = \frac{TP+TN}{TP+TN+FP+FN} \quad (8)$$

Experimental analyzes took an average of 16 seconds per epoch and the training-test success plot for the analyzes is shown in Fig.1. The analysis results obtained from the training are given in Table 3. The confusion matrix is given in Table 4. The success rates of the proposed approach for different epochs are given in Table 5. In addition, information about the total time spent on training the model is given in Table 5. The total training time of the model took 1659 seconds.

Table III. Metric results obtained in experimental analysis

Activity Type	Pre	Rec	F-scr	Overall Acc
Nothing	1.00	1.00	1.00	0.98
Standing still	1.00	1.00	1.00	
Sitting and relaxing	1.00	1.00	1.00	
Lying down	0.99	1.00	1.00	
Walking	0.99	0.99	0.99	
Climbing stairs	0.99	1.00	1.00	
Waist bends forward	1.00	0.98	0.99	
Frontal elevation of arms	0.98	1.00	0.99	
Knees bending	0.99	0.99	0.99	
Cycling	1.00	0.99	1.00	
Jogging	0.97	0.99	0.98	
Running	0.92	0.94	0.93	
Jump front & back	1.00	0.83	0.91	

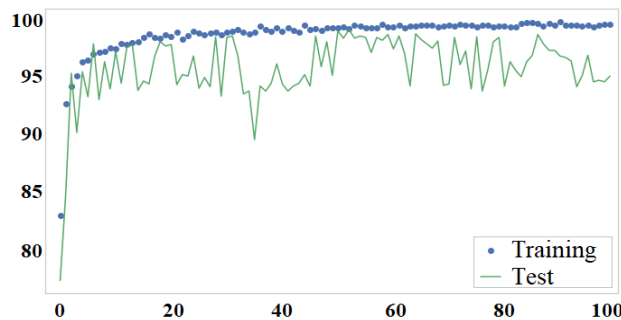


Figure I. Training-test success graphs of the proposed approach

Table IV. The confusion matrix was obtained from the proposed approach. (#1: nothing, #2: standing still, #3: sitting and relaxing, #4: lying down, #5: walking, #6: climbing stairs, #7: waist bends forward, #8: frontal elevation of arms, #9: knees bending, #10: cycling, #11: jogging, #12: running, #13: jump front & back)

	#1	#2	#3	#4	#5	#6	#7	#8	#9	#10	#11	#12	#13
#1	150												
#2		152											
#3			154										
#4				154									
#5				1	149								
#6						106							
#7					2		130	1					
#8								139					
#9								2	145				
#10									1	149			
#11											112	1	
#12											4	60	
#13						1						4	25

Table V. Time - accuracy success chart obtained in the training process.

Epoch	Training Time (sec.)	Training Acc. (%)	Test Acc. (%)
25	442	98.73	96.71
50	846	99.21	97.93
75	1251	99.40	97.14
100	1659	99.64	98.96

As a result of the analysis, the overall accuracy success of the proposed approach in recognizing human activity types was 98.96%. The overall accuracy success obtained from the training data was 99.64%. The proposed approach produced close to 100% results in the analysis of data from the ECG and three types of sensor devices.

## V. Conclusion

Measurement of human activities is carried out using various mechanical/electronic devices. Artificial intelligence-based approaches have become indispensable for a faster and more accurate realization of such systems [30]. In this study, an artificial intelligence-based identification system has been proposed using ECG and sensor measurement data of physical activities. The proposed approach has a structure derived from the integration of the CNN model with the Bi-LSTM model. The contribution of the proposed approach is that it performs more efficient analyzes than traditional methods and can increase success with a hybrid structure. A 98.96% success rate was achieved in the study analyzes. As a result, the contribution of the hybrid approach was observed in the experimental analyzes performed in this study. The results obtained showed that the proposed approach is promising.

In future studies, hybrid models will be designed on video data of real-time human activities. In addition, it will be used in hybrid models in metaheuristic methods that optimize time savings.

## References

- [1] C. Jobanputra, J. Bavishi, and N. Doshi, "Human Activity Recognition: A Survey," *Procedia Comput. Sci.*, vol. 155, pp. 698–703, 2019.
- [2] M.-S. Dao, T.-A. Nguyen-Gia, and V.-C. Mai, "Daily Human Activities Recognition Using Heterogeneous Sensors from Smartphones," *Procedia Comput. Sci.*, vol. 111, pp. 323–328, 2017.
- [3] J. Rose and B. Furneaux, "Innovation Drivers and Outputs for Software Firms: Literature Review and Concept Development," *Adv. Softw. Eng.*, vol. 2016, p. 5126069, 2016.
- [4] R. Vinuesa *et al.*, "The role of artificial intelligence in achieving the Sustainable Development Goals," *Nat. Commun.*, vol. 11, no. 1, p. 233, Jan. 2020.
- [5] E. Prem, "Artificial Intelligence for Innovation in Austria," *Technol. Innov. Manag. Rev.*, vol. 9, no. 12, 2019.
- [6] S. Wan, L. Qi, X. Xu, C. Tong, and Z. Gu, "Deep Learning Models for Real-time Human Activity Recognition with Smartphones," *Mob. Networks Appl.*, vol. 25, no. 2, pp. 743–755, 2020.
- [7] J. Gong, R. Li, H. Yao, X. Kang, and S. Li, "Recognizing Human Daily Activity Using Social Media Sensors and Deep Learning," *Int. J. Environ. Res. Public Health*, vol. 16, no. 20, p. 3955, Oct. 2019.
- [8] N. Jlidi, A. Snoun, T. Bouchrika, O. Jemai, and M. Zaied, "PTLHAR: PoseNet and transfer learning for human activities recognition based on body articulations," in *Proc. SPIE*, 2020, vol. 11433.
- [9] E. Sansano, R. Montoliu, and Ó. Belmonte Fernández, "A study of deep neural networks for human activity recognition," *Comput. Intell.*, vol. 36, no. 3, pp. 1113–1139, Aug. 2020.
- [10] S. Mekruksavanich and A. Jitpattanakul, "Biometric User Identification Based on Human Activity Recognition Using Wearable Sensors: An Experiment Using Deep Learning Models," *Electronics*, vol. 10, no. 3, p. 308, Jan. 2021.
- [11] N. Golestani and M. Moghaddam, "Human activity recognition using magnetic induction-based motion signals and deep recurrent neural networks," *Nat. Commun.*, vol. 11, no. 1, p. 1551, 2020.
- [12] S. Tsokov, M. Lazarova, and A. Aleksieva-Petrova, "Accelerometer-based human activity recognition using 1D convolutional neural network," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 1031, no. 1, p. 12062, 2021.
- [13] G. Jain, "Mobile Health Human Behavior Analysis," Feb-2021. [Online]. Available: <https://www.kaggle.com/gaurav2022/mobile-health>. [Accessed: 30-Mar-2021].
- [14] O. Banos *et al.*, "Design, implementation, and validation of a novel open framework for agile development of mobile health applications.," *Biomed. Eng. Online*, vol. 14 Suppl 2, no. Suppl 2, p. S6, 2015.
- [15] I. N. Yulita, M. I. Fanany, and A. M. Arymuthy, "Bi-directional Long Short-Term Memory using Quantized data of Deep Belief Networks for Sleep Stage Classification," *Procedia Comput. Sci.*, vol. 116, pp. 530–538, 2017.
- [16] C. Zhang, D. Biś, X. Liu, and Z. He, "Biomedical word sense disambiguation with bidirectional long short-term memory and attention-based neural networks," *BMC Bioinformatics*, vol. 20, no. 16, p. 502, 2019.
- [17] R. Yang and Y. Yu, "Artificial Convolutional Neural Network in Object Detection and Semantic Segmentation for Medical Imaging Analysis," *Front. Oncol.*, vol. 11, p. 638182, Mar. 2021.
- [18] S. Asif and K. Amjad, "Automatic COVID-19 Detection from chest radiographic images using Convolutional Neural Network," *medRxiv*, p. 2020.11.08.20228080, Jan. 2020.
- [19] W. S. Ahmed and A. a. A. Karim, "The Impact of Filter Size and Number of Filters on Classification Accuracy in CNN," in *2020 International Conference on Computer Science and Software Engineering (CSASE)*, 2020, pp. 88–93.
- [20] H. J. Jie and P. Wanda, "RunPool: A Dynamic Pooling Layer for Convolution Neural Network," *Int. J. Comput. Intell. Syst.*, vol. 13, no. 1, p. 66, 2020.

- [21] R. Yamashita, M. Nishio, R. K. G. Do, and K. Togashi, "Convolutional neural networks: an overview and application in radiology," *Insights Imaging*, vol. 9, no. 4, pp. 611–629, 2018.
- [22] S. H. S. Basha, S. R. Dubey, V. Pulabaigari, and S. Mukherjee, "Impact of fully connected layers on performance of convolutional neural networks for image classification," *Neurocomputing*, vol. 378, pp. 112–119, 2020.
- [23] H. A. Almurieb and E. S. Bhaya, "SoftMax Neural Best Approximation," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 871, p. 12040, 2020.
- [24] C. Banerjee, T. Mukherjee, and E. Pasilio, "An Empirical Study on Generalizations of the ReLU Activation Function," in *Proceedings of the 2019 ACM Southeast Conference*, 2019, pp. 164–167.
- [25] A. Sawant, M. Bhandari, R. Yadav, R. Yele, and S. Bendale, "Brain Cancer Detection From Mri: a Machine Learning Approach (Tensorflow)," *Int. Res. J. Eng. Technol.*, vol. 05, no. 04, p. 2089, 2018.
- [26] C. Garbin, X. Zhu, and O. Marques, "Dropout vs. batch normalization: an empirical study of their impact to deep learning," *Multimed. Tools Appl.*, vol. 79, no. 19, pp. 12777–12815, 2020.
- [27] F. Demir, A. Şengür, V. Bajaj, and K. Polat, "Towards the classification of heart sounds based on convolutional deep neural network," *Heal. Inf. Sci. Syst.*, vol. 7, no. 1, p. 16, 2019.
- [28] M. Hasnain, M. F. Pasha, I. Ghani, M. Imran, M. Y. Alzahrani, and R. Budiarto, "Evaluating Trust Prediction and Confusion Matrix Measures for Web Services Ranking," *IEEE Access*, vol. 8, pp. 90847–90861, 2020.
- [29] D. Chicco and G. Jurman, "The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation," *BMC Genomics*, vol. 21, no. 1, p. 6, 2020.
- [30] T. Davenport, A. Guha, D. Grewal, and T. Bressgott, "How artificial intelligence will change the future of marketing," *J. Acad. Mark. Sci.*, vol. 48, no. 1, pp. 24–42, 2020.

# Preventing Oscillation of Supply Voltage Due to Resonance Harmonic Frequency of a Power Factor Corrected System

Adel Ridha Othman

Electromechanical Engineering Department  
University of Technology, Baghdad, Iraq  
adel.r.othman@uotechnology.edu.iq

**Abstract-** The variable speed ac and dc drives that are using power electronics are producing high levels of harmonic distortion. The simulated distribution system of this paper consists of a 150 KVA generator supplying a DC drives of 75 KW of type silicon controlled rectifier (SCR). A shunt capacitor bank is connected at a point of connecting the supply with the load which is called point of common coupling (PCC#2) to correct the power factor which is called a power factor correction (PFC) capacitor. The parallel circuit of the PFC capacitor and the inductance of the system has a resonance frequency equals the 11<sup>th</sup> harmonic of the system and that causes an oscillation and distortion in supply voltage. A shunt passive filter, single tuned to the resonance frequency of 11<sup>th</sup> component of harmonics is designed to mitigate the voltage total harmonic distortion (VTHD) to compliance with IEEE Std 519:2014 and preventing the oscillation of the supply voltage.

**Keywords—**Harmonic, resonance, filter, distortion.

## I. Introduction

The aim of studying of harmonic distortion is to calculate the harmonic currents, voltages, and the percentage of the distortion indices in an electrical system and then analyze the situation of resonance and to design filters to mitigate its effect on power system [1]. For reducing harmonics the power converter that must be used must be of operating pulses of 12 and a higher static converter. But due to frequently required maintenance of converters that are of 12-pulse, it is almost using converters of six-pulse instead. The converters of lower number of pulses inject a large values of 5<sup>th</sup>, 7<sup>th</sup> harmonics and the harmonics of orders related to the 12-pulse characteristic harmonics. These are traditionally filtered out by designing a tuned filters to mitigate the low order harmonics such as 5<sup>th</sup>, 7<sup>th</sup>, 11<sup>th</sup> and 13<sup>th</sup>, 17<sup>th</sup> and higher orders [2]. In [3] various passive filters are used to mitigate the harmonic distortion and power factor, in this work single and double tuned filters are designed and investigated for mitigating the harmonics. In [4] the harmonic distortion injected in 20 Kv distribution system is mitigated using passive and active harmonic filters. In [5] a double tuned filter is designed to mitigate the harmonics. In [6] a passive harmonic filters are used to mitigate the source side current harmonics of a rectifier and DC - DC chopper that is feeding a DC motor. In [7] an LC filter is connected at the input of a DC motor to reduce the high frequency harmonics and torque ripples. In [8] the harmonic distortion of industrial sources is

analyzed and a harmonic filters are designed to reduce the effects of these harmonics on the industrial supply. In [9] the harmonic voltage in a large power system that is containing AC- DC converters with a variable large capacity loads is mitigated using passive filter. In [10] different application of passive and active harmonic filters are presented for mitigation of harmonic distortion. In [11] an active filter is used to mitigate the harmonics and a contribution of an easy control method is submitted without needing to transform between power systems frameworks. In [12] a Matlab/Simulink simulation is used to implement an active power filter to select the current harmonics to be mitigated. In [13] a synchronous reference frame theory is used to generate the reference current signal for a voltage source inverter that is controlled by a hysteresis control method and simulating a shunt active power filter based on the synchronous reference frame and hysteresis control method to reduce the harmonic distortion in power system. [14] presents a case study of a hybrid active filter and a passive notch filter to enhance the total harmonic distortion in an industrial power system. [15] presents an active filter to mitigate the generated harmonics from non linear loads of a photovoltaic system connected to the grid. In [16] a Matlab/Simulink software is used to simulate an active filter to enhance the total harmonic distortion in a photovoltaic grid connected system controlled by proportional resonance control method. The effectiveness of the filter is investigated using different load types. [17] presents a Matlab / Simulink modelling of a photovoltaic (PV) grid connected system to study the optimal location of the PV system to the utility grid for minimum harmonic distortion without implementing harmonic filters. In [18] an active filter is used to mitigate the harmonic distortion in wind turbine power plants and the optimal location of the filter is studied. In [19] an active filter is simulated with Matlab/Simulink software to simulate an industrial power system that is using a non linear load of power electronics that inject harmonics to the power system. In [20] a digital simulation of a shunt active filter is used to compensate for harmonics and reactive power. In [21] a single tuned filter is designed using Electromagnetic Transient Analysis Program (ETAP) to mitigate the current harmonics injected by variable speed drives (VSD). In this work the simulated distribution system is composed of a 150KVA generator supplying a DC drives of 75 KW of type of six pulse silicon controlled rectifier (SCR). A parallel capacitor is added at the common coupling point (PCC#2) to correct the power factor (PF). The frequency of resonance of

the parallel circuit of that capacitor and the inductance of the system is happened to be the 11<sup>th</sup> harmonic and that causes an oscillation and distortion in supply voltage. A shunt passive filter single tuned to the 11<sup>th</sup> order harmonic is designed to mitigate the voltage total harmonic distortion (VTHD) to be in compliance with IEEE Std 519:2014 and preventing the oscillation of the supply voltage [22].

## II. The Simulated System:

Equation (1) shows that the resonance frequency can be anticipated and calculated if it is known the value of the short circuit level where the capacitor bank is installed [23].

$$h_r = \sqrt{\frac{KVA_{short\ circuit}}{KVAR_{capacitor\ bank}}} \quad (1)$$

From equation (1) the resonance frequency  $h_r$  can be calculated by knowing the short circuit power level of the system  $kVA_{short\ circuit}$  and the installed capacitor bank reactive power rating  $kVAR_{capacitor\ bank}$ . Figure I shows the one line diagram of the distribution system.

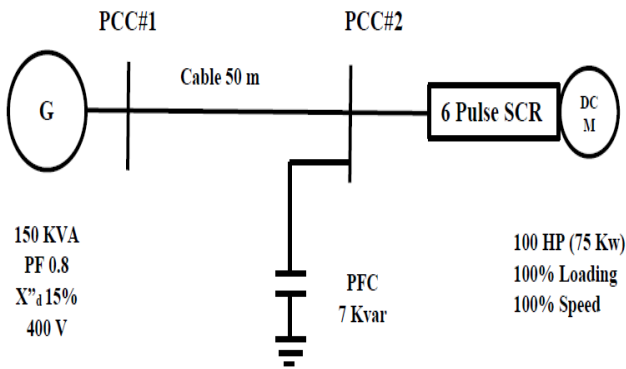


Figure I Simulated distribution system

The voltage, current waveforms and their spectrum of harmonics are shown in figure II.

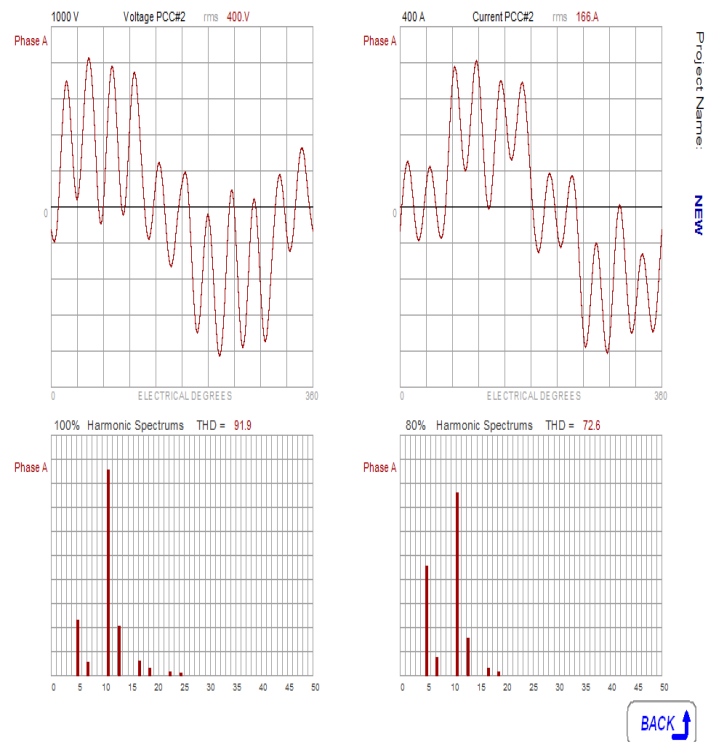


Figure II Voltage and Current Waveforms at Resonance and the Respective Spectrum at PCC#2

The VTHD, the Current Total Harmonic Distortion (ITHD) and the values of their respective harmonics components are shown in table II, as it is seen that they are in compliance with IEEE Std 519:2014.

The calculated parameters of the distribution system are shown in table I

Table I Calculated Parameters at PCC # 2

VTHD (%)	91.6
ITHD (%)	31.5
Load Current $I_L$ (A)	166.2
Short Circuit Current $I_{SC}$ (KA)	1.4
$I_{SC} / I_L$	10.7
Displacement Power Factor (DPF)	0.92 Lag.
True Power Factor (TPF)	0.55
S (KVA)	115.2
Q (KVAR)	96.3
P (KW)	63.2

Table II The Compliance with IEEE Std 519:2014

The Calculated Parameter	Harmonic No.	Calculated Value [%]	IEEE Std 519:2014 Limit [%]	
VTHD	-----	91.6	8.0	FAIL
Maximum Voltage Distortion of Harmonics Components (MVDHC)	11	87.5	5.0	FAIL
ITHD	-----	31.5	5.0	FAIL
Maximum Current Distortion of Harmonics Components (MIDHC) < 11	5	29.5	4.0	FAIL
MIDHC 11 to 16	11	7.3	2.0	FAIL
MIDHC 17 to 22	17	2.8	1.5	FAIL
MIDHC 23 to 34	23	1.5	0.6	FAIL
MIDHC > 35	37	0.5	0.3	FAIL

The capacitor bank current waveform and its harmonics spectrum is shown in figure III.

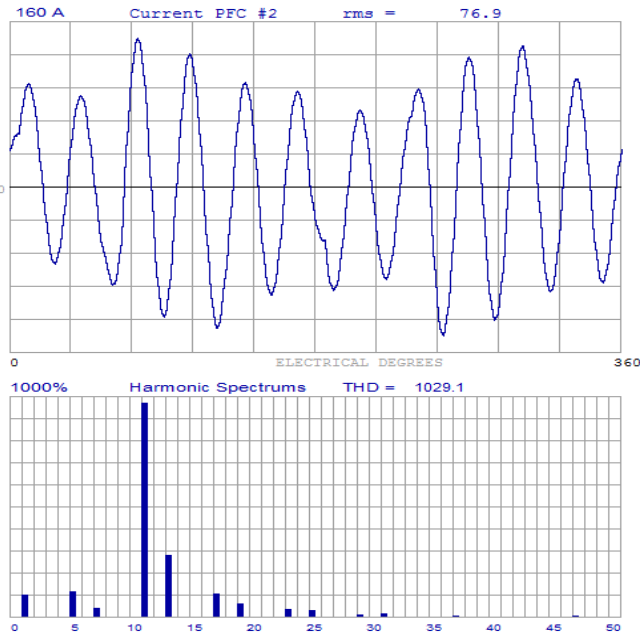


Figure III Capacitor Bank Current Waveform and its Harmonics Spectrum

### III. Design of a Series Passive Single Tuned Filter (SPSTF)

An R-L-C connected in series and tuned to a specified frequency is constitute a series single tuned filter (STF) and that specified frequency is the resonance frequency of the series R-L-C. The 11<sup>th</sup> harmonic component is shown in table II as it is of the highest value among other harmonics. At PCC#2 is connected a filter of type SPSTF tuned to the 11<sup>th</sup> harmonic component. The SPSTF must supply reactive power to compensate the power factor which is calculated to improve the true power factor (TPF) [24]. In this work TPF is improved from 0.55 to 0.9, From table I:-

$$\text{TPF} = 0.55 \quad (\text{True Power Factor})$$

$$\text{Active power} = 63.2 \text{ KW}$$

$$\text{The phase angle of the TPF is } \Phi_{\text{TPF}} = \arccos \text{ TPF} = \arccos 0.55 = 57^\circ \text{ then } \tan \Phi_{\text{TPF}} = 1.5$$

$$\text{And PF is intended to be equal 0.9 then } \Phi_{\text{Required}} = 26^\circ \text{ and } \tan \Phi_{\text{Required}} = 0.5.$$

$$\begin{aligned} \text{The compensation of reactive power} &= \text{load (KW)} [\tan \Phi_{\text{TPF}} - \tan \Phi_{\text{Required}}] \\ &= 63.2 [1.5 - 0.5] \\ &= 63.2 \text{ KVAR} \end{aligned}$$

$$X_C = \frac{V_{LL}^2}{\text{VAR}} = \frac{400^2}{63.2 \times 10^3} = 2.5 \Omega \quad (2)$$

$$X_C = \frac{1}{2\pi f C} \quad (3)$$

$$C = \frac{1}{2\pi f X_C} = \frac{1}{2\pi \times 550 \times 2.5} = 116 \mu\text{f} \quad (4)$$

$$X_C = X_L \quad (5)$$

$$L = \frac{2.5}{2\pi \times 550} = 0.72 \text{ mH} \quad (6)$$

$$R = \frac{X_L}{Q.F} = \frac{2.5}{40} = 0.06 \Omega \quad (7)$$

Where Q.F is the Inductor Quality Factor and supposed to be equals 40

### IV. Simulation Results

The designed and implemented filter of type SPSTF to mitigate the 11<sup>th</sup> harmonic is of capacity 63.2 Kvar. Table III shows the results of the calculated parameters with connected filter and the comparison of these values with the parameters shown in table I it is clearly there is an enhancement of these parameters.

Table III Calculated Parameters at PCC # 2

VTHD (%)	5.3
ITHD (%)	6.3
Load Current $I_L$ (A)	127
Short Circuit Current $I_{sc}$ (KA)	1.4
$I_{sc} / I_L$	11.7
Displacement Power Factor (DPF)	0.95 Lag.
True Power Factor (TPF)	0.95
S (KVA)	88.3
Q (KVAR)	28.9
P (KW)	83.5

Figure IV shows the voltage and current waveforms with their respective harmonic spectrums with connected filter and in comparison with figure II it is clearly that the waveforms and harmonics spectrum are greatly enhanced.

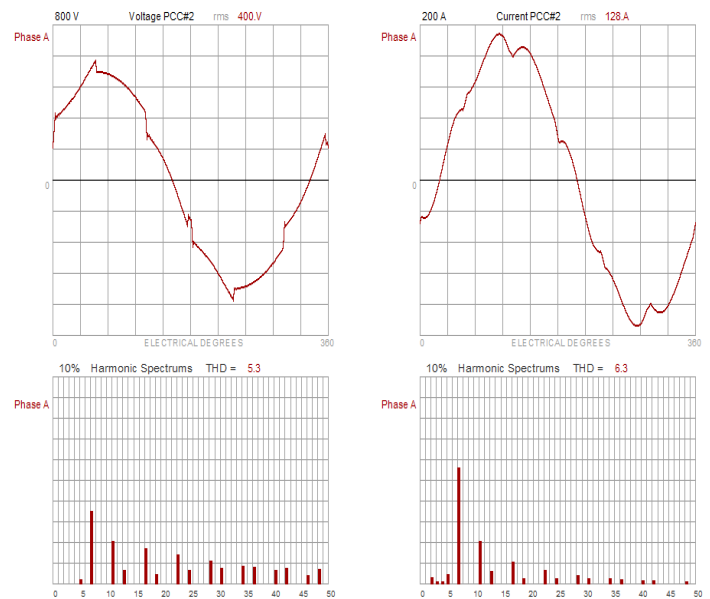


Figure IV Voltage and Current Waveforms at Resonance and the Respective Spectrum at PCC#2

## V. Discussion

Table IV shows that SPSTF of resonance frequency of 11<sup>th</sup> harmonic is managed to mitigate VTHD to be compliance with IEEE Std 519:2014 and the filter has eliminated the resonance, therefore the voltage oscillation is also eliminated as shown in figure IV, the ITHD is still not compliance with or out of the limit of the standard IEEE Std 519: 2014 but it is lowered greatly from 31.5% to 6.3% and also the individual current harmonic components distortion is reduced which is clearly by comparing the values in table II with table IV. Another SPSTF branches can be designed and installed to mitigate the 7<sup>th</sup> harmonic. To mitigate ITHD to be compliance with IEEE Std 519:2014 a high pass filter can be used.

Table IV The Compliance with IEEE Std 519:2014

The Calculated Parameter	Harmonic No.	Calculated Value [%]	IEEE Std 519:2014 Limit [%]	
VTHD	-----	5.3	8.0	PASS
MVDHC	7	3.5	5.0	PASS
ITHD	-----	6.3	5.0	FAIL
MIDHC < 11	7	5.6	4.0	FAIL
MIDHC 11 to 16	11	2.1	2.0	FAIL
MIDHC 17 to 22	17	1.1	1.5	PASS
MIDHC 23 to 34	23	0.7	0.6	FAIL
MIDHC > 35	35	0.3	0.3	PASS

## VI. References

- [1] J. C. Das, "Power System Analysis Short-Circuit Load Flow and Harmonics" Amec, Inc. Atlanta, Georgia
- [2] Arrillaga, J. and Watson, N., "Power Systems Harmonics", 2nd ed., Wiley, New York, 2003.
- [3] S.N. AL. Yousif, M. Z. C. Wanik, A. Mohamed, "Implementation of Different Passive Filter Designs for Harmonic Mitigation" National Power & Energy Conference (PECon) 2004 Proceedings, Kuala Lumpur, Malaysia
- [4] Muhammad Rusli, Muhammad Ihsan, Danang Setiawan, "Single Tuned Harmonic Filter Design As Total Harmonic Distortion Compensator". 23rd International Conference On Electricity Distribution Lyon, 15-18 June 2015
- [5] HE Yi-Hong, SU Heng, "A New Method of Designing Double-Tuned Filter". Proceedings Of The 2nd International Conference On Computer Science And Electronics Engineering (ICCSEE 2013)
- [6] H. Prasad, M. Chilambaran, T.D. Sudhakar, "Application Of Passive Harmonic Filters To Mitigate Source Side Current Harmonics In An Ac - Dc - Dc System", IJRET: International Journal Of Research In Engineering And Technology. Volume: 03 Issue: 01 Jan-2014, Available @ Http://Www.Ijret.Org
- [7] A. Albert Rajan, Dr. S. Vasantharathna, "Harmonics And Torque Ripple Minimization Using L-C Filter For Brushless DC Motors", International Journal Of Recent Trends In Engineering, Vol 2, No. 5, November 2009
- [8] Thet Mon Aye, Soe Win Naing, "Analysis Of Harmonic Reduction By Using Passive Harmonic Filters", International Journal of Scientific Engineering and Technology Research, ISSN 2319-8885 Vol.03, Issue.45 December-2014, Pages:9142-9147
- [9] Byungju Park, Jaehyeong Lee, Hangkyu Yoo, And Gilsoo Jang, "Article Harmonic Mitigation Using Passive Harmonic Filters: Case Study In A Steel Mill Power System", Energies 2021, 14, 2278. <https://doi.org/10.3390/En14082278>
- [10] Lukas Motta, Nicolás Faúndes, "Active / Passive Harmonic Filters: Applications, Challenges & Trends", 978-1-5090-3792-6/16/\$31.00 ©2016 IEEE.
- [11] A. Medina-Rios, and H. A. Ramos-Carranza, "An Active Power Filter in Phase Coordinates for Harmonic Mitigation", Ieee Transactions On Power Delivery, Vol. 22, No. 3, July 2007.
- [12] L. A. Cleary-Balderas, A. Medina-Rios, "Selective Harmonic Current Mitigation with a Shunt Active Power Filter," 978-1-4673-2308-6/12/\$31.00 ©2012 IEEE.
- [13] Kakoli Bhattacharjee, "Harmonic Mitigation by SRF Theory Based Active Power Filter using Adaptive Hysteresis Control", 2014 Power and Energy Systems: Towards Sustainable Energy (PESTSE 2014).
- [14] Henning Tischer, Tomaz Pfeifer, "Hybrid Filter for Dynamic Harmonics Filtering and Reduction of Commutation Notches - A Case Study", 978-1-5090-3792-6/16/\$31.00 ©2016 IEEE.
- [15] Mohamed J. M. A. Rasul, H.V. Khang, Mohan Kolhe, "Harmonic Mitigation of a Grid-connected Photovoltaic System using Shunt Active Filter", 978-1-5386-3246-8/17/\$31.00 © 2017 IEEE.
- [16] Juan C. Colque, Jos´e L. Azcue, Ernesto Ruppert, 'Photovoltaic system grid-connected with active power filter functions for mitigate current harmonics feeding



nonlinear loads”, 2018 13th IEEE International Conference on Industry Applications.

- [17] Mariem Yakoub Yousef, Mohamed Mahmoud Ismail, Said Mohamed El masry, “The Effect of Grid Connected Photovoltaic Location and Penetration level on Total Harmonic Distortion”, 2018 Twentieth International Middle East Power Systems Conference (MEPCON), Cairo University, Egypt, 978-1-5386-6654-8/18/\$31.00©2018 IEEE.
- [18] Daphne Schwanz, Math Bollen, Anders Larsson, Łukasz Hubert Kocewiak, “Harmonic Mitigation in Wind Power Plants: active filter solutions”, 978-1-5090-3792-6/16/\$31.00 ©2016 IEEE.
- [19] R. Sheba Rani, C. Srinivasa Rao, M. Vijaya Kumar, “Analysis of Active Power Filter for Harmonic Mitigation in Distribution System”, International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT) – 2016, 978-1-4673-9939-5/16/\$31.00 ©2016 IEEE.
- [20] Narayan G. Apte, Vishram N. Bapat, Amruta N. Jog, “A Shunt Active Filter for Reactive Power Compensation and Harmonic Mitigation”, The 7th International Conference on Power Electronics October 22-26, 2007 / EXCO, Daegu, Korea.
- [21] Mohamed Awadalla, Mohamed Orner, Ahmed Mohamed, “Single-tuned Filter Design for Harmonic Mitigation and Optimization with Capacitor Banks,” International Conference on Computing, Control, Networking, Electronics and Embedded Systems Engineering, 978- 1-4673-7869-7/ 15/\$31.00 ©20 15 IEEE.
- [22] Tony Hoevenaars, P. Eng., Kurt Ledoux, P. Eng. And Matt Colosino, “Interpreting IEEE Std 519 And Meeting Its Harmonic Limits In VFD Application”, Copyright Material IEEE, Paper No. PCIC-2003-15.
- [23] Francisco C. De La Rosa, “Harmonics And Power Systems” Distribution Control Systems, Inc. Hazelwood, Missouri, U.S.A. 2006
- [24] Seema P. Diwan, Dr. H. P. Inamdar, and Dr. A. P. Vaidya, “Simulation Studies of Shunt Harmonic Filters: Six Pulse Rectifier Load – Power Factor Improvement and Harmonic Control” Proc. of Int. Conf. on Advances in Electrical & Electronics 2010, ACEE.

# Arranging Spaces for Next Pandemic by inspiring from Qajar period

Parastoo Pourvahidi  
Department of Architecture, Faculty of  
Fine Arts, Design and Architecture,  
Cyprus International University, Via  
Mersin 10, 99258 Nicosia, Turkey  
ppourvahidi@ciu.edu.tr

**Abstract**—this research has been initially commenced by pursuing the Qajar period residential building's situation if it was in pandemic period. The result demonstrated the interesting entity about the organization of spaces which stimulated this research to consider it in contemporary building in Iran. Although today's word situation is totally different by facing cumulative population and lack of land. Though comparing the immense land with building that has more than 4 bedrooms to building that has 2 bedrooms is excruciating. Coronavirus changed the life style of people all around the world. In the past if building without balcony has opening toward greenery area was a valuable option. But during pandemic, necessity for having semi-open spaces such as balcony become more imperative. The pandemic period taught people, open and semi-open spaces is as much vital as close space if you are during lock down period. Subsequently, this study attempts to check the total depth value (TDn) of the spaces in Qajar period to find out the solution for today's apartment building that unfortunately, open spaces is much smaller or invaluable than close spaces. Residents with low and medium level in economy could not have a prime to buy a new apartment, hence, renovating the space arrangement can be the surpassed and faster option. This research used the depthmap-x as a method for the purpose of suggesting simple solution, try to encourage the residents to re-evaluate the worth of life and try to have more connection with fresh area as their ancestor did in the past. Space syntax as a method used in this investigation to aid the research for discovering the shortest distance or total depth value (TDn) and integration value (i) of each space after repositioning the spaces.

**Keywords:** *pandemic, space arrangement, depthmap, Qajar period, contemporary building*

## I. Introduction

Pandemic of Coronavirus (COVID-19) in almost all the cities around the world has caused the restriction for public movement and as a result people locked down in their home [13]. Lock down and movement restriction origin the fundamental changes in convention of urban space and also in socializing behavior. In this sense Tamborrino (2020) contends that “despite global alarm about the spread of the novel coronavirus, this desertification has been a gradual process, as people showed resistance to the interruption of the use of social spaces” [15]. Restriction was not just for social distancing such as work space control, facility closure, travel restriction or curfews, similarly it has effect on the organization of space in resident's way of life. In that manner for decreasing the spread of Coronavirus, recommended to people all around the world banned from being in public. Since, the spread of virus slaughtered over 250,000 people

beside of infecting over 3 million people in 185 countries global [16].

All these restriction for controlling the spread of virus all around the world cause the noteworthy altering in the life style of people. For instance, spaces were used by people such as public and private for responding of changing life style after Coronavirus cause to have delicate redesign process. Subjects such as proximity, centrality and clustering concertation and density are revealing the indirect process, which design scholars like planner, urban designer and architect mentioned before about these negative impressions in quality life of people. However, because of socioeconomic purposes all this impression has been overlooked by social performers [16]. Balcony can perform the best semi-open spaces for having social relation with neighbors and decreasing the stress of solitude. Since, during pandemic, balcony (semi-open or open space) supports the residents to have social relation with neighbors with distinguishing type of reserved pleasantness. For instance, in the city of Beirut, balcony is confronted to the street and obtainable company without stresses of closeness [17]. In this concern, World health origination proposing important factor in preventing the virus that cause Coronavirus from spreading indoors. One of them was concerning for having natural ventilation in indoor spaces such as Opening window. In addition, increasing air filtration is another recommendation. Significantly, cumulative the total airflow supply to occupied spaces is another issue which can reduce the spread of virus in close spaces [10]. Furthermore, Kong, in 2017 claimed that semi-open spaces increase the content of the user from 80 % to near 100 %, since it has a latent to offer microenvironment. It means that semi-open space is personalized to the individual preference of the user [9]. All of these concerns cause to this research to compare the modern apartment building with the traditional building by using depthmap-x analysis method. In this case what was happening if pandemic period was happen in Qajar period? And what user can do for preparation for another pandemic with fewer cost and faster method.

## II. Qajar and contemporary period

Iranian traditional houses designed base on privacy and decency. Thus, the spaces are organized in a way not to permit for instant expose to the house. The arrangement of the space lead user first to enter to the house, move around and later they could have an access to varied spaces in a house. Yard in traditional houses is located in the heart of the building, thus users should pass through the passage to have an access to the yard Spaces such as Shahneshtin, five-door and a space which is used for parties named as a big hall are called as public zones. Jelosara where is front room is the closest space to the

yard. Spaces such as access room, house pool and interior yard are the one which is used by family members. In these spaces even close relative and family members meet up for doing activates. Furthermore, spaces like basement, the two-door room and back room are considered as private zone in Iranian traditional houses. These spaces had been used for relation, study, sleep and chat therefore they are positioned at the utmost part of the house [5]. Figure I represent the arrangement of the spaces in traditional houses.

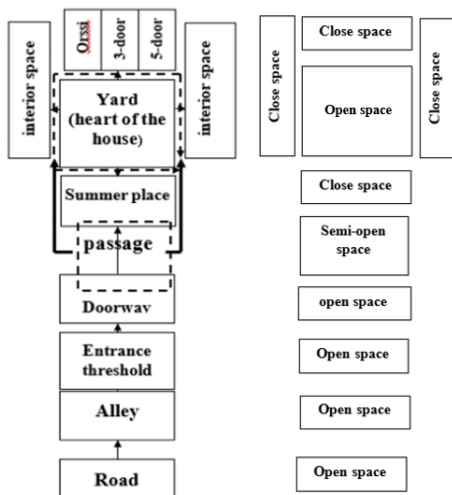


Figure I: Space arrangement and hierarchy in Iranian traditional houses [5].

Based on public and private zones of each spaces, they can be used for dwelling separated or in combination, also all these spaces are providing comfortable condition for the users. In traditional houses there is different kind of spaces such as the open spaces without roof (the yard), with roof without wall (the porch) and closed spaces with roof. Comfortable living condition had been provided for the users in these traditional houses by combing or connecting all three spaces together. Respecting the life style by being flexible, accomplish the user's desires. Creating the larger spaces needs flexibility. Since flexibility is the competence of the spaces to combine with the nearby areas. Windows, Orssi or curtains can be used as an element for connecting the private area to the public spaces through the middle spaces. For instance, house pool and the room, the Shahneshin and Gooshvareh, the two-door room and the Shahneshin, three-door room and the back room and the yard and the Shahneshin can create the new area by connecting together. Considering both interior and exterior and degree of neglecting can assistance to comprehend these spaces, which is the main principle for the space arrangement in the Iranian traditional houses [5].

Moreover, Safarkhani (2016), stated about the different size and location of the balcony in Iranian contemporary houses. In these houses, small percentage by considering the whole plan is belong to the balcony's size. In these kinds of houses, living room and hall are combined together. Furthermore, toilet and bathroom are also separated colored as, blue and purple [12].

By comparing the space arrangement in Qajar and contemporary period, it is noticeable that during Qajar period after passing from street, resident was entering to the open

space such as courtyard, after courtyard users enter to the doorway which was again semi-open space. However, today's residential building plan composed of completely close area with small percentage of semi-open spaces (balcony). Balcony in most of the contemporary building is small hence, residents always change the function of it to the storage rather a sitting and enjoying the comfortable space. The lack of space, increasing population and price of the land based on the close spaces is the result of all these changes during passing years. But pandemic period changed this routine and alarm the resident about the value of open and semi-open spaces in house for communication with neighbors and freshen the air and spending lockdown in open space without scaring about virus. Furthermore, in figure II arrangement of the spaces for contemporary building plan allow the users to content the necessity of having semi-open spaces such as balcony.

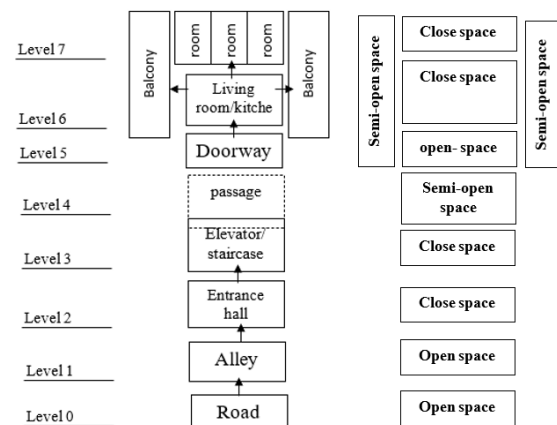


Figure II. Proposing space arrangement for contemporary building based on Qajar period.

It is not senseless to notify that inhabitant should change the house for being ready for the next pandemic. But, renovating and reconstructing the interior space can reasonable. For instance, in most of the cases in Iranian apartment building balcony is located next to kitchen with limited square meter. Or if the inhabitant is lucky, there is more than one balcony in one of the bedrooms as well. But because of restricted accessibility and limited square meter, this kind of semi-open spaces are useless. So as a solution this research attempt to rearrange the organization of spaces in case to have a direct accessibility to the balcony from living room. In addition, using flexible partition wall can also give opportunities to the residents to extend the living room toward the balcony during spring and summer time. Furthermore, by closing the partition wall could limit the access during winter time and balcony can be usable as a sunspace for warming the living room. Another prosper of having balcony with the extension o living room is that during pandemic resident could invite guest in open space by obeying the social distance. In the following analyzing of the Agraph regarding the cases has been presented both before and after renovation.

### III. ANALYSIS AND DISCUSSION

#### A. connectivity Method

As Bahreini in 2014 mentioned that architect has been always looking for understanding the invisible structures and organization. Since these factors are located behind the

architectural form and phenomena. Space syntax defined as one of the methods which can explore the space morphology [3]. Space syntax method had a purpose to notice the social relationship between the spaces [6]. In another point of view, space syntax attempt to detect the reason for the independence of each space and try to state each space based on the position of it [7]. There is different kind of software which is applicable for space syntax method, this research used the depthmap X for the analysis.

Depthmap X is analysis software in different space for representing the spatial network. Alasdair Turner (Space syntax group) developed this software. The purpose of this software is to create the map which has been shown the spatial element and relationship between them [4]. Hence, Depthmap X is one of the methods which is used in this research.

Depth map analysis manifest the values for TDn, MDn, I, CV and RA for each space. Also, Ostwald mentioned that for comparing to building with each other, outcome prerequisite to be stabilized in term of relative depth, which is called as Relative Asymmetry (RA). This result could be between 0 to 1[11]. Segregation specifies by high value of Relative Asymmetry and integration specify by low value of Relative Asymmetry [8].

RA is the reflection of its relative isolation for a carrier space. Afterwards, the i can be calculated by the shared of that node. Table below represent the formula of the relationship between RA and I value (table I) [11].

Table I: RA and i formulas [11].

RA	Relative asymmetry	A measure of how deep a system is relative to a symmetrical of balanced model for the same system. RA is calculated by the following formula: $RA = 2(MD-1)/K-2$
i	Integration value	A measure of the degree of integration (relative centrality of spaces) in a system. i is the reciprocal of RA, thus $i = 1/RA$

Before, starting the analysis with depthmap X, this research generally analyzed the spaces in traditional building plan during Qajar period. The analysis has been clarified such as:

The proportion of open and semi-open spaces (blue color) and closes spaces (yellow color) demonstrated that open spaces is as much important as close spaces during Qajar period. In addition, during pandemic, they could spend more time outside, even they could invite the people, and there is enough space for social activities (fig.III).



Figure III: Traditional building plan in Qajar period (a), indicating open (blue) and close spaces (yellow).

These dash lines demonstrated the combination of living room with courtyard. However, the interesting thing about these spaces is the partition wall, which was flexible.

Pandemic period demonstrates the fact how open plan create problem for inhabitant while one of them should do online work and the children should be online in school but with no partition wall concertation would be impossible however in this Qajar period's plan by having flexible partition wall separating the spaces is possible and beneficial during pandemic (fig.IV).

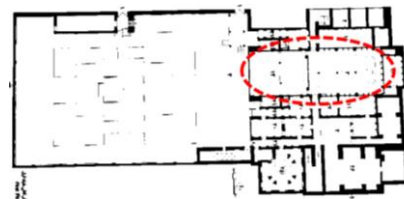


Figure IV: Dash lined represent the flexible space in Qajar period.

One of the first spaces after courtyard is bathroom. It is located before living room and bedrooms, which gives the opportunities to the inhabitant to hygiene they hand before entering to the building, which is one of the huge problems in contemporary houses during pandemic (fig.V).

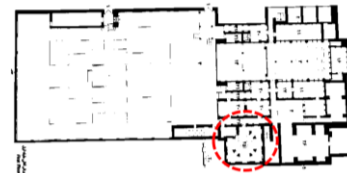


Figure V: Dash lined represent the bathroom space in Qajar period.

There is special space, which they call it removing shoes spaces. This space is open space in case to remove the smell and is the second spaces after connecting corridor to the living room.

These kinds of space are essential during Coronavirus disease's time in case to remove all the dirty shoes outside of the living space in and staying in open space to become cleanness with sunlight (fig.VI).



Figure VI: Dash lined represent the Antechamber

Designing the corridor (yellow color) between each space in another positive fact in the plan of Qajar period. Entering directly to the space with shoes can cause a lot of problem during pandemic. Though, if each space separated with connecting corridor can be helpful for inhabitant to stay sanitation from Coronavirus disease (fig.VII).



Figure VII: connecting corridor placed beyond each spaces

Depth map analysis in table II illustrates the axial analysis that blue line (low connectivity) demonstrates the sanitation area and red line (high connectivity) the dirty spaces in both Qjara and contemporary buildings. Generally, drawing the graph in case studies aid this research to manifest the connectivity's of each space similarly understanding the hygiene of the spaces is conceivable. Hence, this research just analyzing the spaces based on integration value.

Table II: Analyzing of plan in Qajar period throughout pandemic

Depth map analysis	<p>Blue line, which has a function such as room, are the safest and cleanest spaces. Even the Living room has the color light blue which again demonstrated the less integration.</p>
Agraph analysis	<p>0-entrance, 1-entrance,2-courtyard, 3-Antechamber (midway space through which the closed spaces are accessed), 4- antechamber, 5- antechamber, 6- dinning room, 7-room,8-room,9-room, 10-room, 11- bathroom, 12-corridor, 13-corridor, 14-courtyrad, 15- Seh-dari (room with three large contiguous window opening to the courtyard), 16- reception hall (large formal room usually assigned to entertaining guest), 17- reception hall, 18-backyard, 19-staircase, 20- kitchen, 21-Panj-dari (large room with five large contiguous window opening to the courtyard, 22-storage, 23- kitchen storage, 24- room, 25-corridor.</p>

		TDn	MDn	RA	i	CV
Result	0 0	99	3.96	0.24	4.05	0.25
	1 1	99	3.96	0.24	4.05	0.25
	2 2	75	3.00	0.16	6.00	2.58
	3 3	88	3.52	0.21	4.76	1.08
	4 4	64	2.56	0.13	7.69	1.95
	5 5	97	3.88	0.24	4.16	0.83
	6 6	79	3.16	0.18	5.55	1.33
	7 7	99	3.96	0.24	4.05	1.00
	8 8	83	3.32	0.19	5.17	0.75
	9 9	72	2.88	0.15	6.38	0.53
	10 10	110	4.40	0.28	3.52	1.33
	11 11	88	3.52	0.21	4.76	0.25
	12 12	59	2.36	0.11	8.82	2.33
	13 13	134	5.36	0.36	2.75	0.50
	14 14	94	3.76	0.23	4.34	2.83
	15 15	75	3.00	0.16	6.00	0.95
	16 16	96	3.84	0.23	4.22	0.58
	17 17	118	4.72	0.31	3.22	0.25
	18 18	73	2.92	0.16	6.25	2.20
	19 19	83	3.32	0.19	5.17	0.20
	20 20	95	3.80	0.23	4.28	1.25
	21 21	118	4.72	0.31	3.22	0.25
	22 22	119	4.76	0.31	3.19	0.50
	23 23	119	4.76	0.31	3.19	0.50

Alitajer in 2016 claim that in Iranian traditional houses, the highest percentage of connectivity was in yard and corridor which are positioned between inside (*Andarooni*) and outside (*birooni*). Furthermore, the minimum percentage belong to the spaces such as bedrooms, toilet and spaces, which is, just belong to the women (*Pastoo*). However, in modern and contemporary houses in Iran, the connectivity situation is totally different. For instance, the highest connectivity is in corridor and living room. The lowest connectivity in modern houses is in bedrooms, kitchen, bathroom, toilet and corner of balcony (Alitajer & Molavi Nojumi, 2016). Therefore, this research attempt to find out the highest and lowest connectivity in three contemporary cases in Iran, for understanding the better relation between the spaces. Since, the lowest connectivity in balcony, toilet and kitchen is not contest with the impression of this research for preparing the houses for the next pandemic. In the following this research attempt to detect the connectivity's beyond the contemporary building in Iran (table III and table IV).

Table III. Analysis of contemporary building case No.A with Agraph method

	Plan/ depth map	Graph
Depth map analysis	<p>[12]</p>	<p>0. Entrance, 1. WC,</p>

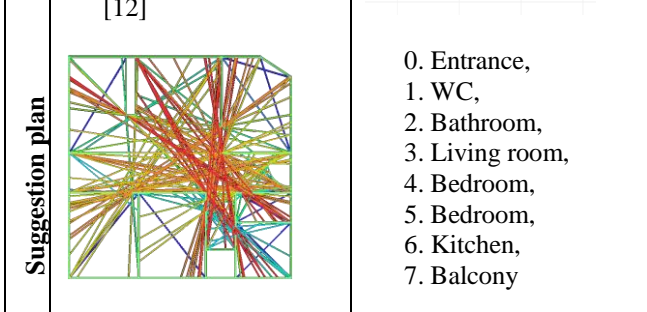
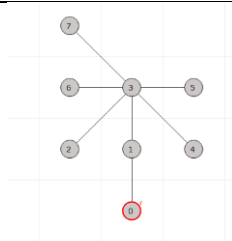


2. Bathroom,
3. Living room,
4. Bedroom,
5. Bedroom,
6. Kitchen,
7. Balcony

	TDn	MDn	RA	i	CV	
0	0	19	2.71	0.57	1.75	0.50
1	1	13	1.85	0.28	3.50	1.20
2	2	15	2.14	0.38	2.62	0.20
3	3	9	1.28	0.09	10.50	4.00
4	4	15	2.14	0.38	2.62	0.20
5	5	15	2.14	0.38	2.62	0.20
6	6	13	1.85	0.28	3.50	1.20
7	7	19	2.71	0.57	1.75	0.50
Min	9.00	1.28	0.09	1.75	0.20	
Mean	14.75	2.10	0.36	3.60	1.00	
Max	19.00	2.71	0.57	10.50	4.00	

\_summary.html

The highest i represent the dirty space which is living room in this case. The lowest i represent the cleanest space which is balcony and entrance space. Therefore there is something wrong in organization of space in contemporary building. That is why balcony should be the cleanest space?



0. Entrance,
1. WC,
2. Bathroom,
3. Living room,
4. Bedroom,
5. Bedroom,
6. Kitchen,
7. Balcony

	TDn	MDn	RA	i	CV	
0	0	18	2.57	0.52	1.90	0.50
1	1	12	1.71	0.23	4.20	1.16
2	2	14	2.00	0.33	3.00	0.16
3	3	8	1.14	0.04	21.00	5.50
4	4	14	2.00	0.33	3.00	0.16
5	5	14	2.00	0.33	3.00	0.16
6	6	14	2.00	0.33	3.00	0.16
7	7	14	2.00	0.33	3.00	0.16
Min	8.00	1.14	0.04	1.90	0.16	
Mean	13.50	1.92	0.30	5.26	1.00	
Max	18.00	2.57	0.52	21.00	5.50	

\_summary.html

After rearranging the spaces, the i value of balcony changes to the same value like bedrooms. Although the i value of living room still is so high but at least it is

connected to the balcony (semi-open spaces) for refreshing the air in that space.

Table IV. Analysis of contemporary building case No.B with Agraph method

	Plan/ depth map	Graph
Depth map analysis (connectivity)		
		<ol style="list-style-type: none"> <li>0. Entrance,</li> <li>1. Living room,</li> <li>2. Kitchen,</li> <li>3. Balcony,</li> <li>4. WC,</li> <li>5. Bathroom,</li> <li>6. Bedroom,</li> <li>7. Bedroom,</li> <li>8. Balcony.</li> </ol>

	TDn	MDn	RA	i	CV	
0	0	17	2.12	0.32	3.11	0.16
1	1	10	1.25	0.07	14.00	5.00
2	2	15	1.87	0.25	4.00	1.16
3	3	22	2.75	0.50	2.00	0.50
4	4	17	2.12	0.32	3.11	0.16
5	5	17	2.12	0.32	3.11	0.16
6	6	17	2.12	0.32	3.11	0.16
7	7	15	1.87	0.25	4.00	1.16
8	8	22	2.75	0.50	2.00	0.50
Min	10.00	1.25	0.07	2.00	0.16	
Mean	16.88	2.11	0.31	4.27	1.00	
Max	22.00	2.75	0.50	14.00	5.00	

\_summary.html

In this case study, highest i (dirty space) value is belong to living room and the lowest I (clean space) value is belong to balcony.

	Plan/ depth map	Graph
Suggestion plan		
		<ol style="list-style-type: none"> <li>0. Entrance,</li> <li>1. Living room,</li> <li>2. Kitchen,</li> <li>3. balcony,</li> <li>4. WC</li> <li>5. Bathroom,</li> <li>6. Bedroom,</li> <li>7. Bedroom,</li> <li>8. Balcony,</li> <li>9. WC,</li> <li>10. Balcony.</li> </ol>

		TDn	MDn	RA	i	CV
0	0	20	2,00	0,22	4,50	0,11
1	1	11	1,10	0,02	45,00	8,50
2	2	20	2,00	0,22	4,50	0,11
3	3	20	2,00	0,22	4,50	0,11
4	4	20	2,00	0,22	4,50	0,11
5	5	20	2,00	0,22	4,50	0,11
6	6	20	2,00	0,22	4,50	0,11
7	7	18	1,80	0,17	5,62	1,11
8	8	27	2,70	0,37	2,64	0,50
9	9	20	2,00	0,22	4,50	0,11
10	10	20	2,00	0,22	4,50	0,11
	Min	11,00	1,10	0,02	2,64	0,11
	Mean	19,63	1,96	0,21	8,11	1,00
	Max	27,00	2,70	0,37	45,00	8,50
_summary.html						
After the rearranging the space, balcony's I value increase same as the bedroom which is convenient.						

In case number one (A) (table III), case number 2 (B) (table VI) before renovation, connectivity is highest in the living room which is  $i=21$  and  $i=14$  and lowest in the balcony (A.  $i=1.75$ ) and (B.  $i=2$ ) and bedroom (A.  $i=2.62$ ) (A.  $i=3.11$ ). Also, after renovation, connectivity is highest in the living room (A.45) and (B.  $i=21$ ) and balcony (A.  $i=3$ ) and (B.  $i=4.5$ ). Lowest in two bedroom (A. $i=3$ ) and (B.  $i=4.5$ ). Before, rearranging the spaces, balcony (open spaces) integration value is less than other spaces, even it is less than entrance. However, it is the best space during pandemic that people can have social distancing. Balcony in some of the buildings located after kitchen area. It means the access to this space it should be from kitchen. WC is located next to the entrance door which is good for hygiene prior to entering to the main space. Kitchen can be considered as balcony that is next to the entrance and kitchen could be places partially in living room. Partition between living room and balcony can be flexible wall in case to extend the living room during winter time likewise extend the balcony during summer time. Generally, the result of all the cases based on Agraph method analysis demonstrated the interesting values in TDn and integration for balconies formerly and after renovating plan. TDn value of balcony before renovation has the same with the entrance gate, but after extending balcony toward living room the total depth value decreases noticeable. In addition, the integration value before renovating the plan was the lowest integration value but after renovation plan it is increased which is the content value for integration of balcony. Commonly, extending the balcony toward living represent the nourishing outcome which this research was expected. Pandemic period teach architect designing the beautiful close space is not adequate for pleasing the user since having open spaces is much needed as close spaces.

Also based on the analysis, there is a decline of RA value in both cases after renovation. This means that the segregation between the spaces (nodes) convert to integration due to diminishing the value of RA. Generally, decreasing the RA value in spaces like balcony (semi-open) after renovation represent the more integration of this spaces which is satisfying. Subsequently, the mean RA value in Qajar period is 0.22 which by comparing to the contemporary houses was lower than them. Now that RA value of case A is 0.36 and after renovation is 0.30. RA value of case B is 0.31 and after renovation is RA=0.21. Hence, decline of the RA values after re-organizing the spaces develop content.

Moreover, in traditional house in Qajar period, the lowest connectivity is detected in bedrooms and store room. However, the highest connectivity is detected in in yards and corridor. Furthermore, the kitchen, bathroom represent the low connectivity as well. Nevertheless, in contemporary building the lowest connectivity detected the bathroom and toilet and the highest connectivity is detected on living room. Therefore, based on all the analysis these are the recommendation rule which can be supplementary by the municipality's rule for the modern building, which resident can be prepared for the next pandemic:

#### Rules and regulation for next pandemic:

- Integration value should be high for living room and balcony
- Balcony should not have low integration.
- The wall of the balcony should be flexible in order to be extended toward living room based on the appeal of the user.
- Living room must be connected to semi-open spaces (balcony)
- Having open space such as courtyard and balcony is the must in each residential building.

There should be balance in arrangement of semi-open, open and close spaces. (Pandemic period evident the fact, people prefer during the say spend some time in the balcony to just see the outside for having peace of mind. Nonetheless unfortunately, the square meter of the balcony than close space is less than half in modern building).

Living room with Partition wall (inhabitant could still keep the open plan but by adding partition glass or wall, they could have opportunities to separate the area whenever the online class of children is starting or they should do online working and it needs concertation.

Roof function (most of the apartment building have immense area in roof which is useless, CONID-19 clarify the fact people need social interaction so may be roof can be the space which inhabitant who lives in apartment building can gather there since it is open space)

Positioning the WC at first next to the entrance door (washing hands was the first and primitive way to stay safe during pandemic period, thus locating the WC next to the entrance door could be the first space resident can enter to become sanitization ad after that entering to the main space)

Positing corridor before entering to the private spaces such as room (connectivity spaces such as corridor create the chances to not entering the private spaces directly, therefore, it can stay disinfectant)

Entrance hall (this space can be useful for removing shoes at first in the entrance section of the house, which unfortunately obliterated form the space organization in the modern building. Nevertheless, this space can be semi-open for better ventilation or by using mechanical ventilating the space.)

#### IV. Conclusion

Cocid-19 open signify that designing the plan of apartment houses should change since the user requirement is changes as well. Pandemic period states another desires of the user, such as need for socializing in open space without having a fear of interacting with virus. Afterwards, people try to start

socializing with each other from small balcony in case to nourishing their feeling during lock down. Thus, balcony become the spot in all building for the place which user can have connection with outside and also with their neighbors. What if the balcony has the potential for becoming extended from balcony for inviting the friends without fearing of dispersing the virus? Renovation and adding one flexible glass wall in living room, which can have an access toward open spaces for having, better indoor ventilation and entering the solar radiation, can be the rapid and worthwhile solution for enduring another pandemic. Also, this research recommended to add the mandatory rule in municipality that doorway and accessible spaces such as passage should be open or semi-open spaces. Furthermore, the percentage of the open space (balcony) toward the close spaces should have specific dimension for averting havoc for the next pandemic period.

### References

- [1] SAFARKHANI, M. (2016). N PARTIAL FULFILLMENT OF THE REQUIREMENTS for THE DEGREE OF MASTER OF ARCHITECTURE IN ARCHITECTURE. THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES of MIDDLE EAST TECHNICAL UNIVERSITY.
- [2] Alitajer, S., & Molavi Nojumi, G. (2016). Privacy at home: Analysis of behavioral patterns in the spatial configuration of traditional and modern houses in the city of Hamedan based on the notion of space syntax. *Frontiers of Architectural Research*, 341-352.
- [3] Bahrainy, H., & Taghabon, S. (2015). Deficiency of the space syntax method as an urban design tool, designing traditional urban space and the need for some supplementary methods. *space Ontology International journal*, 1-18.
- [4] depthmapX development team. . (2017). Retrieved from [depthmapX \(Version 0.6.0\)](https://github.com/SpaceGroupUCL/depthmapX/) : <https://github.com/SpaceGroupUCL/depthmapX/>
- [5] Hasan Zolfagharzadeh, H., Jafariha, R., & Delzende, A. (2017). Different Ways of Organizing Space Based on the Architectural Models of Traditional Houses: A New Approach to Designing Modern Houses: (Case Study: Qazvin's Traditional Houses). *Space Ontology International Journal*, 6(4), 17 - 31.
- [6] Hillier, B. (1996). *Space is the machine, A configurational theory of architecture*. Cambridge : Cambridge University Press.
- [7] Hillier, B. (2005). The art of place and the science of space. *World architecture* 11(185), 96-102.
- [8] Khadiga, O. M., & Mamoun , S. (n.d.). the space syntax methodology: fits and misfits. *Art and compout*, 189-204.
- [9] Kong, M. (2017). *Semi-Open Space and Micro-Semi-Open Space and Micro-Environmental Control for Improvement for Improving*. Dissertations - ALL. 810: Syracuse University.
- [10] organization, w. h. (2020, July 29 ). Coronavirus disease (COVID-19): Ventilation and air conditioning in public spaces and buildings. Retrieved from Q&A Detail: <https://www.who.int/news-room/q-a-detail/coronavirus-disease-covid-19-ventilation-and-air-conditioning-in-public-spaces-and-buildings>
- [11] Ostwald , M. J. (2011). The mathematics of spatial configuration: revisiting revising and critiquing justified plan graph theory. *Nexus Network journal*, 445-470.
- [12] Safarkhani, M. (2016). N PARTIAL FULFILLMENT OF THE REQUIREMENTS for THE DEGREE OF MASTER OF ARCHITECTURE IN ARCHITECTURE. THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES of MIDDLE EAST TECHNICAL UNIVERSITY.
- [13] Sandford, A. (2020). Sandford, A. (2020). Coronavirus: Half of humanity now on lockdown as 90 countries call for confinement. *Euronews*.
- [14] Shahid Beheshti university, c. h. (1998). *Ganjnameh*. Tehran: Faculty of architecture and urban planning documentation and research center.
- [15] Tamborrino, R. (2020). Here's how locking down Italy's urban spaces has changed daily life. [weforum.org](http://weforum.org).
- [16] Ülkeryıldız, E., Vural, D., & Yıldız, D. (6-8 May 2020). Transformation of Public and Private Spaces: Instrumentality of Restrictions on the. 3rd International Conference of Contemporary Affairs in Architecture and Urbanism (ICCAUA-2020).
- [17] Zacka, B. (2020, May 11). An ode to the humble balcony, in times of the pandemic. Retrieved from DTNEXT: [www.DTNEXT.IN/NEWS](http://www.DTNEXT.IN/NEWS)



# Comparison of Sentiment-Lexicon-based and Topic-based Sentiment Analysis Approaches on E-Commerce Datasets

Adeola O. Opesade  
Department of Data and Information  
Science, Faculty of Multidisciplinary  
Studies, University of Ibadan, Nigeria  
morecrown@gmail.com

**Abstract**— Discovering underlying sentiment in a user’s textual data is a complex task; nevertheless, human beings have been intuitive enough to interpret the tone of a piece of writing. The hugeness of online reviews, due to advancements in internet-based applications, has however, made the need for computer-based models highly imperative for sentiment analysis of texts and speeches. Many of the existing studies have examined the performance of either of the two main sentiment analysis approaches, symbolic and topic-based approaches. The present study investigated comparatively, the performances of Liu Hu sentiment-lexicon implementation and bag of words topic-based approaches. The study revealed amongst others that sentiment analysis, like other data mining tasks, is an experimental science. It recommends that analysts could compare the performances of symbolic and topic-based approaches in their sentiment classification endeavors when deciding on the most precise technique to adopt.

**Keywords**— Sentiment analysis, Amazon customer reviews, Konga customer reviews, Liu-Hu sentiment-lexicon, Textual data mining

## I. Introduction

Advancements in internet-based applications have fuelled the availability of huge volumes of personalised reviews on the Web [1]. This user-generated data, mostly unstructured, usually carry elements of user opinions and sentiments about goods, services, events and experiences in the online or offline environments [2]. These reviews are becoming so increasingly important that many people now consult them as sources of information to aid their understanding, planning and decision making processes [1]. Businesses have also adopted online reviews as part of their criteria for quality assessments [3]. Little wonder then that learning customers’ emotional inclinations through online reviews is becoming more crucial in the present Information Age [4].

Discovering underlying sentiments, based on user’s textual data, is not a trivial task. This is especially due to the different intricacies associated with language, such as contextual differences, language implications and sentiment indistinctness of certain words. Also, some writers could be sarcastic while some others may not express specific sentiment markers in their writing [3]. Despite these complexities, human beings have been found to be passably intuitive in interpreting the tone of a piece of writing [5]. The massiveness of online reviews has however, made human beings to rely on computer-based models in identifying the

polarity of sentiments expressed by writers through the process known as sentiment analysis [6], [2].

Sentiment analysis, an intellectual process of extracting user’s feelings and emotions contained in a piece of writing or speech, is a language processing task that uses a computational approach to identify opinionated content and categorize it as neutral, positive or negative [7], [2]. It is one of the fields of Natural Language Processing (NLP) and data mining that has gained popularity in the recent years [1] [3]. A lot of research work is being carried out in the field of sentiment analysis [1]. These studies employed different data mining and Natural language methods, mainly classified as symbolic and machine learning approaches.

The symbolic approach, also known as rule-based classification entails reliance on sentiment-lexicon to find the polarity of each word in a review; if the number of words tagged positive is greater than that tagged negative, it is concluded that the writer’s sentiment is positive, otherwise, it is said to be negative [8]. It is therefore said to be a great knowledge-based classification that might lack generality due to possibility of its closeness to specific linguistic and operational fields [4]. Machine learning approach also regarded as a topic-based text classification approach is a general resolution, independent on any special fields. In this approach, the reviews are represented by different features, followed by any text classification algorithm [4] [8].

In order to identify better alternatives in sentiment analysis, a number of studies have used rule-based approaches, combined with some text processing procedures and machine learning algorithms, to investigate the performances of sentiment analysis tasks [3] [7] [9]. A number of studies have also investigated the performances of sentiment analysis tasks based on machine learning approach by using bag of words, n-grams and POS-tagged feature selection techniques, combined with series of text processing procedures and machine learning classifiers [6] [10] [11]. It could be observed that most of these previous studies have investigated sentiment polarity of textual data either from sentiment-lexicon or topic-based approach. They have therefore, reported the relative performances of sentiment classification tasks, based either on one or the other sentiment analysis approach. While the approaches employed might

suffice for the researcher’s purposes of investigation, there still remains a dearth of information on the relative performances of sentiment lexicon-based and topic-based polarity approaches on the same dataset. This is particularly important because of variability in real datasets and the fact that universality of learning algorithms has been said to be a mere fantasy [12]. How will sentiment-lexicon and topic-based approaches perform on a dataset? Answer to this question will help to investigate the universality of sentiment analysis approach further.

The present study, in order to provide an answer to this question, investigated the performances of both approaches on two different e-commerce datasets. Two datasets were examined for the purpose of triangulation. To achieve the objective of the study, the following research questions were specifically examined.

1. How do machine assigned sentiment-lexicon based score compare with human assigned sentiment label?
2. How do sentiment-lexicon-based classification schemes compare with topic-based classification schemes in each dataset?
3. Which classification scheme is the best for each dataset?

## II. Research Methods and Materials

The method of research adopted by the present study is the textual data mining, with supervised machine learning technique.

### A. Data Collection

Two web based electronic commerce datasets were used in the present study. The datasets are:

1. Dataset on Amazon: This dataset was collected from [13]. The dataset was created and uploaded by [14], authors of 'From Group to Individual Labels using Deep Features'. It comprises of one thousand reviews of products labelled with positive or negative sentiment. The authors of the dataset selected review sentences that have a clearly positive or negative connotation in equal proportion. The dataset therefore, contains 500 positive and negative reviews each.

2. Dataset on Konga: Konga is one of the most popular e-commerce site in Nigeria. Users’ comments on services received from Konga was collected by the author of the present study. Data collection was carried out on the 17<sup>th</sup> of April 2020 on Twitter, using #kongacustomer as search term on Orange data mining tool Twitter API. Tweets were first read and labelled based on the sentiment inclination expressed therein. These sentiment polarity labels were positive, negative and neutral. However, all neutral tweets were removed from the dataset in order to make the dataset to conform to sentiment polarity format of the collected Amazon dataset.

## B. Methodology

The methodologies employed for data analytic procedures are the Liu Hu and bag of words implementations in Orange data mining tool. Liu Hu is a lexicon-based sentiment analysis technique that computes a single normalized score of sentiment in the text (negative score for negative sentiment, positive for positive, 0 is neutral). The technique was used to carry out sentiment-lexicon based scoring of each review. Bag of words technique was used to extract a number of most frequent content words in each corpus. With this technique, words (excluding stop words) were successively extracted to contain 1000, 500, 200, 100, 50, 20 and then 10 most frequent words from each dataset. Orange data mining tool was also used to carry out text pre-processing (transformation, tokenization, removal of stop words and extraction of word features) and machine learning classification experiments.

## C. Experimental Setup

An experiment was carried out, using six machine learning algorithms in Orange data mining tool. The machine learning algorithms used were K-Nearest Neighbor (KNN), Tree, Support Vector Machine (SVM), Neural Network (NN), Naïve Bayes (NB) and Logistic Regression (LR). The experiment was carried out to determine the performances of these machine learning classification schemes based on the following feature sets:

- a. Sentiment-lexicon-based machine assigned score.
- b. Topic-based bag of word vectors (1000, 500, 200, 100, 50, 20 and 10 most frequent content words).

Ten fold cross validation was used to evaluate the models' performances based on Classification Accuracy (CA) and F-Measure (F1).

## III. Results and Discussion

Research Question 1: How do machine assigned sentiment-lexicon-based scores compare with human assigned sentiment labels?

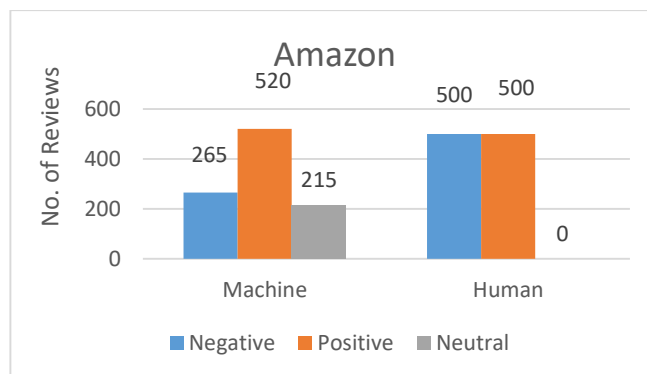


Figure 1a. Sentiment labels of Amazon dataset reviews

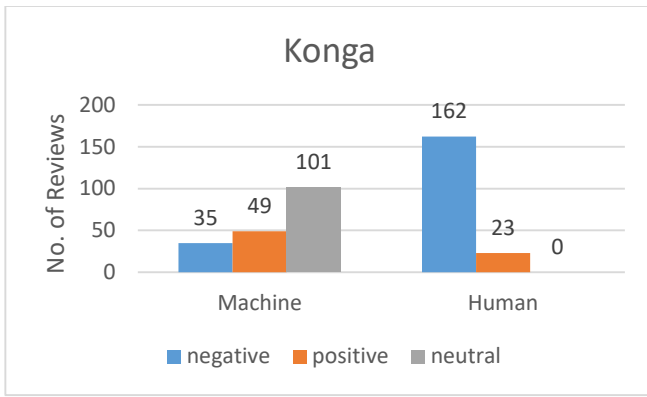


Figure 1b. Sentiment labels of Konga dataset reviews

Whereas none of the two datasets contain neutral reviews, the machine sentiment-lexicon scoring algorithm scored some of the reviews as neutral. For example, as shown in Figure 1a, the Amazon dataset originally contains 500 positive and negative reviews each, as labelled by human identifier. The machine sentimental algorithm however, scored 520 reviews as positive, 265 as negative and 215 as neutral. Also as shown in Figure 1b, while the original konga dataset contains 23 positive and 162 negative reviews as labelled based on human identification, machine sentimental algorithm identified 49 reviews as positive, 35 as negative and 101 as neutral. This shows that machine scoring, based on sentiment-lexicon algorithm, finds it more challenging than human agents to determine sentiment polarity of the reviews. This corroborates the assertion that human beings are fairly more intuitive than machine when it comes to interpreting the tone of a piece of writing [5].

Research Question 2: How do sentiment-based classification schemes compare with topic-based classification schemes in each dataset?

The results of sentiment-lexicon and bag of words machine learning classification of Amazon data with the six classification algorithms, KNN, Tree, SVM, NN, NB and LR

are as presented in Figure 2a and Figure 2b respectively, while the results of the same experimental task on Konga data are as presented in Figure 3a and Figure 3b respectively.

In the case of Amazon dataset (Figures 2a and 2b), the sentiment-lexicon-based approach outperformed the bag of words sentiment polarity machine learning classification across the six classification algorithms. In the case of Konga dataset (Figures 3a and 3b), sentiment-lexicon-based approach outperformed the bag of words topic-based sentiment polarity approach in two classifiers (Tree and SVM). The 200 word and 500 word topic-based feature sets outperformed sentiment-lexicon-based approach for KNN, all examined bag of words feature sets other than 50 and 100 words outperformed sentiment-lexicon-based approach in NN, fifty (50) word feature outperformed the sentiment-lexicon-based approach in NB while the sentiment-lexicon-based approach was outperformed by all bag of words feature set options in LR.

It could be observed that the behavior of the two data sets are not exactly the same. While Amazon data was better determined by the sentiment dictionary based approach, the same did not hold for Konga data. This divergence might possibly be traced to the assertion of [4] that since the sentiment dictionary based approach is a great knowledge-based classification approach, it lacks generality, and that of [15] who reported that non frequently mentioned features are often not detected by this knowledge-based sentiment analysis approach. It could possibly mean that most of the sentiment clues in Amazon dataset are better represented in the lexicon database than those in Konga dataset. Nevertheless, the result further buttresses the assertion that real datasets vary and idea of a universal approach is just a fantasy [12]. Thus, Sentiment Analysis, like other data mining tasks, is an experimental science.

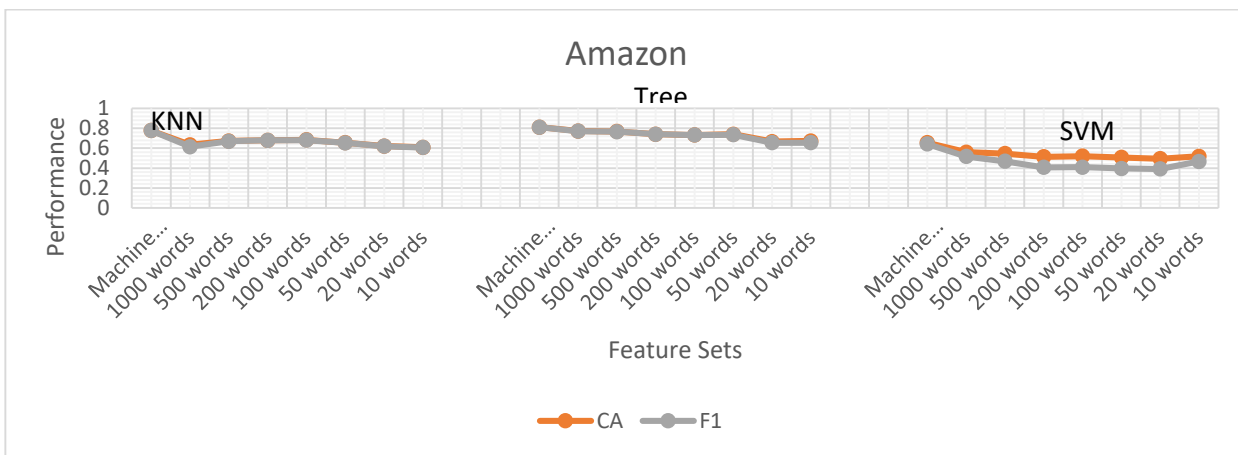


Figure 2a. Performance of sentiment-lexicon-based analysis classification on Amazon data

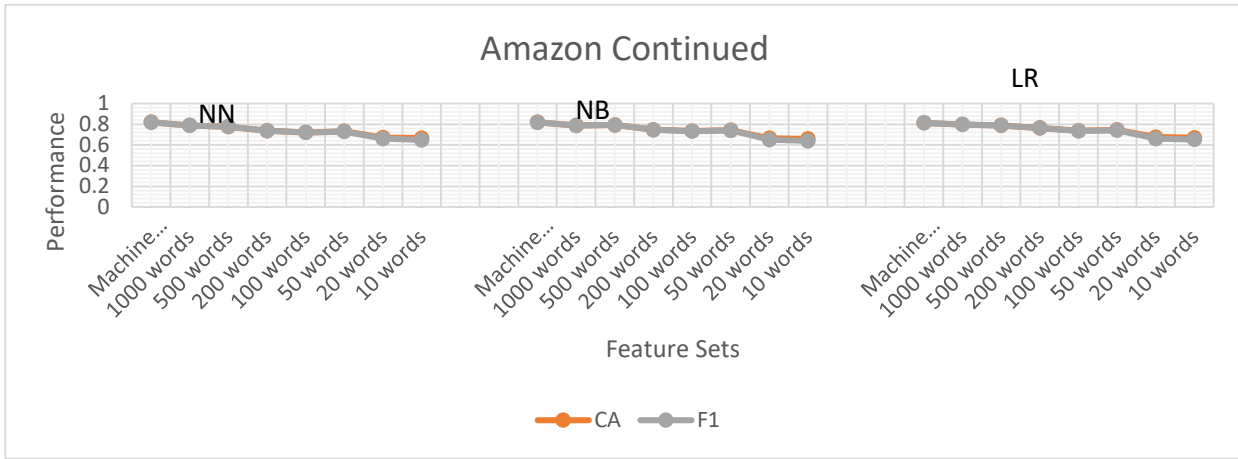


Figure IIb. Performance of sentiment analysis classification on Amazon data continued

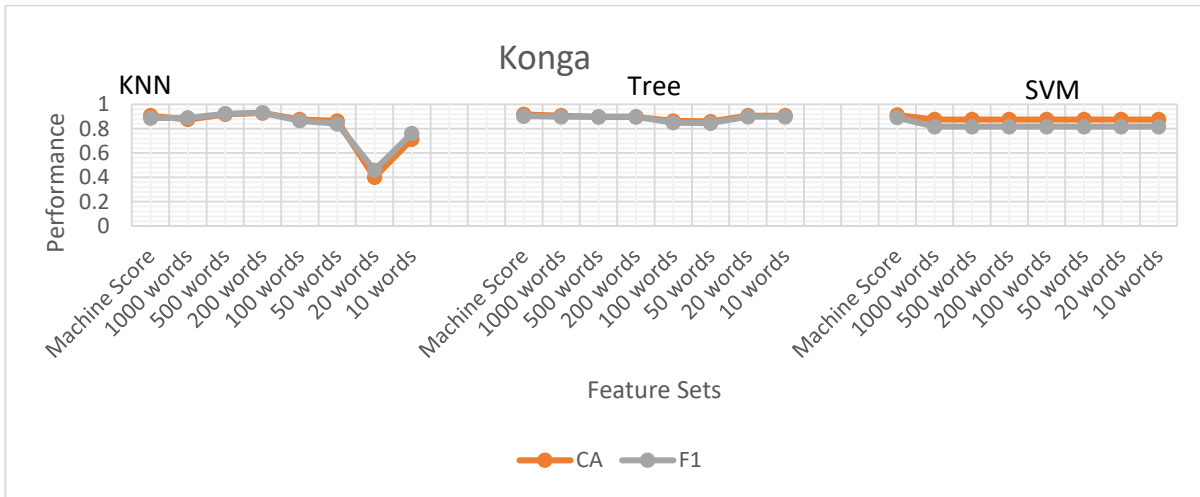


Figure IIIa. Performance of sentiment analysis classification on Konga data

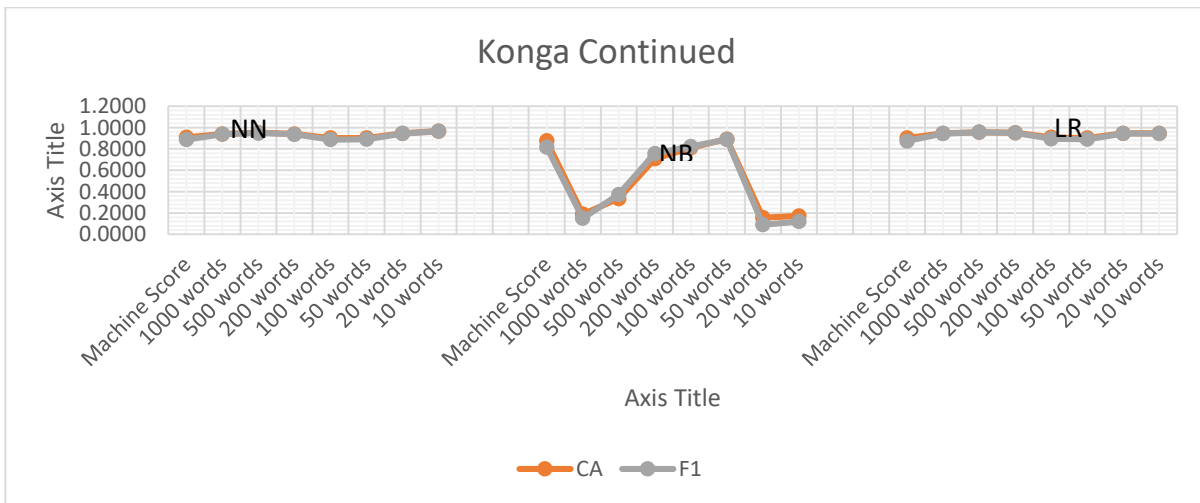


Figure IIIb. Performance of sentiment analysis classification on Konga data continued

Research Question 3: Which classification scheme is the best for each dataset?

The best performing feature set in each of the six classifiers for the two datasets are as shown in Tables I and II.

Table I. Amazon best performing classification schemes

Algorithm	Feature set	CA	F1
KNN	Sentiment	0.7774	0.7771
Tree	Sentiment	0.8106	0.8104
SVM	Sentiment	0.6544	0.6429
NN	Sentiment	<b>0.8203</b>	<b>0.8203</b>
NB	Sentiment	0.8188	0.8188
LR	Sentiment	0.8135	0.8134

Table II. Konga best performing classification schemes

Algorithm	Feature set	CA	F1
KNN	200 words	0.9297	0.9316
Tree	Sentiment	0.9189	0.9045
SVM	Sentiment	0.9135	0.8925
NN	10 words	<b>0.9676</b>	<b>0.9654</b>
NB	50 words	0.8919	0.8874
LR	500 words	0.9568	0.9559

As shown in Tables I and II, sentiment-lexicon approach and ten word topic-based feature sets, based on Neural Network,

provided the highest classification accuracy in the Amazon and Konga data sets respectively. The emergence of NN as the best classifier in the two datasets is at variance with some of the previous studies that reported SVM [6], NB [7], or Stochastic Gradient Descent (SGD) [3] as the best classifiers in their studies. It could however, be observed that none of the mentioned previous studies included NN as a classifier in their experiments. The difference in the most performing feature sets in the present study, and the emergence of NN as the best classifier contrary to reported cases, further buttress the fact that the concept of universal learner is an idealistic fantasy [12].

#### IV. Conclusion

Based on the findings of this study, it could be concluded that rule-based sentiment analysis approach, particularly the Liu Hu dictionary implementation used in the present study has not yet assumed the human status in sentiment polarity approach. The study also concludes that sentiment analysis, like other data mining tasks is an experimental science and analysts could compare the performances of symbolic and topic-based approaches in their sentiment classification endeavor before deciding on the most precise technique.

The sentiment-lexicon results in the present study were however, based on Liu Hu dictionary implementation alone. Also the topic-based approach was based on bag of words lexical features alone. Performances of other methods of each of the two techniques could be investigated in further studies.

#### References

- [1] M. . Devika, S. C., and A. Ganesh, "Sentiment Analysis: A Comparative Study On Different Approaches," *Procedia Comput. Sci.*, vol. 87, pp. 44–49, 2016, doi: 10.1016/j.procs.2016.05.124.
- [2] V. K. Singh, R. Piryani, A. Uddin, and P. Waila, "Sentiment Analysis of Textual Reviews," no. December 2015, 2013, doi: 10.1109/KST.2013.6512800.
- [3] A. Salinca, "Business reviews classification using sentiment analysis," no. September, 2015, doi: 10.1109/SYNASC.2015.46.
- [4] L. Zhang, K. Hua, H. Wang, G. Qian, and L. Zhang, "Sentiment Analysis on Reviews of Mobile Users," *Procedia - Procedia Comput. Sci.*, vol. 34, pp. 458–465, 2014, doi: 10.1016/j.procs.2014.07.013.
- [5] K. Bannister, "Sentiment Analysis: How Does It Work? Why Should We Use It? | Brandwatch," *Brandwatch.Com*. 2018, [Online]. Available: <https://www.brandwatch.com/blog/understanding-sentiment-analysis/>.
- [6] T. U. Haque and F. M. Shah, "Sentiment Analysis on Large Scale Amazon Product Reviews," no. June 2019, 2018, doi: 10.1109/ICIRD.2018.8376299.
- [7] R. S. Jagdale, V. S. Shirsat, and S. N. Deshmukh, "Sentiment Analysis on Product Reviews Using Machine Learning Techniques," in *Proceeding of CISC 2017 Sentiment Analysis on Product Reviews Using Machine Learning Techniques*, 2019, no. January, doi: 10.1007/978-981-13-0617-4.
- [8] S. Mukherjee and P. Bhattacharyya, "Feature Specific Sentiment Analysis for Product Reviews," pp. 1–12.
- [9] Y. Xu, X. Wu, and Q. Wang, "Sentiment Analysis of Yelp's Ratings Based on Text Reviews," *2015 17th Int. Symp. Symb. Numer. Algorithms Sci. Comput.*, 2015.
- [10] Z. Singla, S. Randhawa, and S. Jain, "Sentiment Analysis of Customer Product Reviews Using Machine Learning," in *2017 International Conference on Intelligent Computing and Control (I2C2)*, 2017, no. June, doi: 10.1109/I2C2.2017.8321910.

- [11] W. Kasper and M. Vela, "Sentiment Analysis for Hotel Reviews," in *Proceedings of the Computational Linguistics-Applications Conference*, 2011, vol. 231527, pp. 45–52.
  
- [12] I. H. Witten and E. Frank, *Data Mining: Practical Machine Learning Tools and Techniques 2nd edition*, vol. 5, no. 1. 2005.
  
- [13] Kaggle, "Sentiment Labelled Sentences Data Set," 2021. <https://www.kaggle.com/marklvl/sentiment-labelled-sentences-data-set>.
  
- [14] D. Kotzias, M. Denil, N. De Freitas, and P. Smyth, "From group to individual labels using deep features," *Proc. ACM SIGKDD Int. Conf. Knowl. Discov. Data Min.*, vol. 2015-Augus, pp. 597–606, 2015, doi: 10.1145/2783258.2783380.
  
- [15] E. Guzman and W. Maalej, "How Do Users Like This Feature ? A Fine Grained Sentiment Analysis of App Reviews," in *2014 IEEE 22nd International Requirements Engineering Conference (RE)*, 2014, pp. 153–162, doi: 10.1109/RE.2014.6912257.

# Autism detection from facial images using deep learning methods

Abdulazeez Mousa  
Department of Software Engineering  
Firat university  
Elazig, Turkey  
abdulazizmousa93@gmail.com

Fatih Özyurt  
Department of Software Engineering  
Firat university  
Elazig, Turkey  
fatihozyurt@firat.edu.tr

Shivan H. M. Mohammed  
Computer Science  
Duhok University  
Duhok, Iraq  
shivan.cs@uod.ac

**Abstract**— Autism spectrum disorder (ASD) refers to a collection of behavioral and developmental issues and difficulties. The cognitive, communication, and play skills of a child with autism spectrum disorder are all affected. The description "spectrum" in autism spectrum disorder refers to the fact that each child is special and has their own set of characteristics different from other children. These come together to give him a unique social bond as well as his own understanding of his own actions. Medical image classification is a significant research field that is gaining traction among researchers and clinicians alike to detect and diagnosis diseases. It addresses the issue of medical diagnosis, experiment purposes and analysis in the field of medicine. To address and resolve these issues, a number of data mining-based medical imaging modalities and applications have been proposed and developed. But also to understand and learn about how diseases develop in patients, to help doctors in early diagnosis of pathology. Not only is achieving good accuracy in classifying medical images is the main purpose alone. We used pre-trained Convolutional Neural Networks and transfer learning in this study. These CNN architectures are used to train the network and to classifying medical images. The experiment results of this study show that the proposed model can detect Autism Spectrum Disorder (ASD), with the best accuracy of 95.75 percent achieved using the MobileNet model with transfer learning. The architectures that tested in this study are ready to be tested with additional data and can be used to prescreen Individuals with ASD. The use of a deep learning methods for feature selection and classification in this study could greatly support future autism studies.

**Keywords**—Autism, deep learning, Convolutional Neural Networks, Classification, Transfer Learning

## I. INTRODUCTION

Autism, also known as autism spectrum disorder (ASD), is a developmental disorder that affects a person's interaction, speech, and social behavior like (Strange gestures, such as swinging hands and rotating, are repeated. Since autistic people can't stay in one place for long periods of time, we find that the patient are always moving. and the autistic person's gestures are disorganized and spontaneous. Can't connect well with others and speaks with a peculiar accent. Autistic people are very sensitive to light and sound[1]. In addition, they are unaware of other people's emotions and thoughts. When compared to other children, they are more violent. They frequently have outbursts of anger. The pain response and sensation are sluggish and mild. Depending on the disposition of the child, who may be slow to learn or have a high level of intelligence.) and other disorders such as depression, anxiety, and attention deficit hyperactivity disorder are common in these children [2]. Early diagnosis in infancy can help autistic

children develop their communication and social skills, as well as their quality of life. Estimating the total ASD requires the expertise of an ASD specialist; however, early diagnosis is crucial for controlling and treating this disease [3]. However, in rural areas and isolated villages, health care and facilities are unavailable. Several methods for determining whether or not a child has autism spectrum disorder have been established. These methods are extremely useful for early diagnosis of ASD and evaluating the efficacy of the ASD protocol [4]. In latest years, deep learning (DL) has advanced rapidly. Image processing, computer-aided diagnosis, image recognition, image fusion, image registration, image segmentation, and other fields have benefited from deep learning techniques, By detection and analysis of medical images Deep learning techniques extracts features from medical images accurately and efficiently reflect the information. Deep learning (DL) helps physicists (medical practitioners) and doctors detect and predict disease risk more accurately and quickly, allowing them to avoid disease before it develops [5]. One of the most dilemmas (challenges) in image recognition field is the classification of medical images. Classification of medical images with the aim of analyzing and categorizing them (classification) into several different groups to assist physicists (medical practitioners) and doctors in diagnosing diseases or performing research on them [6]. These techniques enhance the abilities of doctors and researchers to understand that how to analyze the generic variations which will lead to disease. Deep learning algorithms such as convolutional neural networks (CNNs) to detect and classify images. The classification of medical images is divided into several stages. Firstly getting medical images and uploading them to the model. And then extraction of essential features from the acquired input image is performed. The features are then used to create models that identify the image data set in the second phase of medical image classification. The final stage is the production, which is a classified image and a report based on image analysis that reflects the outcome [7].

The rest of this article is structured as follows as: we presents the previous related work about autism detection and image classification using deep learning and transfer learning techniques in Section II, The details of the dataset used in this study and description of the CNN architectures that implemented on our dataset provided in the section III, The experimental results of the CNN architectures and transfer learning techniques that performed on the dataset and the show that the best results obtained of all models performed is MobileNet model presents in section IV, Finally in section V presents the conclusion of proposed framework.

## II. LITERATURE REVIEW

Beary, M., Hadsell, A., Messersmith, R., & Hosseini, M. P. proposed "Diagnosis of Autism in Children using Facial Analysis and Deep Learning" In this paper, the authors proposed a deep learning model to classify children if they are autistic or non-autistic (healthy). Autistic patients in general have distinct facial abnormalities that are part of the disease discovery that allows researchers to know if a person has autism or not. Therefore, the researchers in this study used the MobileNet deep learning model with two dense layers to extract the features of the face and classify the people according to the image. In this study, the model was trained and tested on 3014 images (90% for training, 10% for testing). The results of this study was 94.6% accuracy in classifying [8]. Mythili, M. S., & Shanavas, A. M. have written "A study on Autism spectrum disorders using classification techniques" In this study, the authors focused on extracting data from methodologies to study the performance of students with autism. By mining data with the tasks it provides that can be used to study the performance of students with autism. In this article the authors use algorithms for Data mining by task classification at the level of students. The authors used Support Vector Machines (SVM), Artificial Neural Networks (ANN), and Fuzzy logic as machine learning methods in this paper. The algorithms are extremely helpful in dealing with the prediction level of autism students [9]. Ali, N.A., Syafeeza, A.R., Jaafar, A., & Alif, M have written "Autism spectrum disorder classification on electroencephalogram signal using deep learning algorithm " In this paper, the authors use deep learning algorithms via electroencephalogram (EEG) to detect the different patterns between normal children and children with autism. The process of classifying normal children and children with autism by using deep learning algorithms and using the Multilayer Perceptron (MLP) model in this article is by means of a database that contains brain signals for pattern recognition, using the Multilayer Perceptron (MLP) model to extract the features for the classification process [10]. Rad, N. M., Kia, S. M., Zarbo, C., van Laarhoven, T., Jurman, G., Venuti, P., Marchiori, E., & Furlanello, C. have written "Deep learning for automatic stereotypical motor movement detection using wearable sensors in autism spectrum disorders" in this paper the authors proposed deep learning application to aid automatic stereotypical motor movements (SMM) detection with multi-axis Inertial Measurement Units(IMUs). In the study from raw data convolutional neural network (CNN) have been to learn a discriminative feature space, this method is used. To model the temporal patterns in a series of multi-axis signals, combine the long short-term memory (LSTM) with CNN architectures. Furthermore, it was demonstrated how the convolutional neural network (CNN) can be used with transfer learning technique to improve the detection and analyzation rate on longitudinal data. From the results of this paper they shows that Handcrafted features are outperformed by feature learning, the detection rate improves when using LSTM to learn the temporal dynamic of signals, particularly when the training data is distorted. Detectors with an ensemble of LSTMs are more accurate and stable. In longitudinal settings, parameter transfer learning is beneficial. These results represent a significant step forward in detecting SMM in real-time scenarios [11].

## III. METHODOLOGY

### A. Dataset

Early detection of autism is very important for the development of the affected child, so the development of these classifications based on a set of data related to it is a reason for early detection of autism and this enhances the adaptation of the patient to normal life. The dataset can be used in two different ways, both of which are popular for deep learning tasks. The dataset is divided into Training, Test, and Validation sets, which is a normal procedure. Train is the title of the training set. That is contains of two subdirectories known as (Autistic & Non\_Autistic) [12, 13]. There are 1667 facial pictures of children with autism in the Autistic subdirectory in 224 X 224 X 3, jpg format. And also there are 1667 photos of children who do not have Autism in the Non\_Autistic subdirectory, all of which are in the 224 X 224 X 3 jpg format. The unified directory is the second way the data is made accessible. There are two subdirectories in this directory: Autistic & Non\_Autistic. It is a collection of files from the train, test, and correct directories that have been combined into a single set. The combined data can then be partitioned into user-defined train, test, and validation sets.

Some images of Autism dataset used in our study shown below in Figure 1.



Figure 1: some images of (a) Autistic and (b) Non-Autistic Children.

### B. Transfer Learning and CNN architectures

Due to the emergence of modern and high-performing machine learning systems, image classification has been more important in the research field. Artificial neural networks(ANNs) have progressed, and deep learning architectures such as the convolutional neural network(CNN) have emerged. The application of multi - class image classification and recognition of objects belonging to multiple categories has been triggered by depending on artificial neural networks(ANN). In terms of efficiency and complexity, a new machine learning(ML) algorithm has an advantage over older ones[14]. In the deep learning field, transfer learning is a research problem. It extracts the knowledge collected from one problem and applies it to a separate but related issue. For example, knowledge acquired while learning to recognize a disease may apply when trying to identify another disease[15, 16]. For example, we apply knowledge gained during the identification of cancer to malaria. Transfer learning is a deep learning technique that involves training a neural network



model on a problem that is close to the one being solved. One of the advantages of deep learning is that it shortens the training time on the deep learning model and this way we can build on previous knowledge rather than starting from scratch. Transfer learning is commonly expressed in computer vision by the use of pre-trained models. A pre-trained model is one that has been trained on a broad benchmark dataset to solve a problem that is close to the one we are working on. As a result of the high computational cost of training these models [17, 4].

Convolutional neural networks (ConvNets or CNNs) are one of the most common types of neural networks used to recognize and classify images. CNNs are commonly used in areas such as object detection, face recognition, and so on. In classification of images Convolutional neural network (CNN) takes an image as input, process (extract features from it) it and categorize it [18]. (Like: Autistic, Non-Autistic). An input image is seen by computers as an array of pixels, with the number of pixels varying depending on the image resolution. Depending on the resolution of the picture.  $h \times w \times d$  ( $h$  = Height,  $w$  = Width,  $d$  = Dimension) will be shown. For example, a  $6 \times 6 \times 3$  array of RGB matrix (3 refers to RGB values) and a  $4 \times 4 \times 1$  array of grayscale matrix image.

When re-purposing a pre-trained model for our own needs, it should remove the original classifier, then add a new classifier that is appropriate for our needs, and finally fine-tune the model using one of three strategies [19].

**Train the Entire Model** - Apply the pre-trained model's applied syntax to your dataset and train it. Start with values from a pre-trained model instead of random weights.

**Feature extraction** (freezing convolutional neural network (CNN) model base): The pre-trained basic form on which to train a new classifier. In other words, training of the fully connected layer alone. With leaving the weights of the convolution layers without changing anything from them.

**Fine-tuning**: In this strategy the Original convolution layer weights are used as starting points. In addition to a fully connected classifier one or more convolution layers that are retrained. Also, just to fit a new problem, the unsecured convolution layers are adjusted.

There are many pre-trained CNN networks that only require sets of data that consist of training and test data in their own input layer and have the ability to transfer learning. These networks are different in their structures in terms of the internal layers and the technologies they use [20]. In this study we have chosen five pre-trained CNN architectures to implement on our autism dataset. The CNN architectures (Densenet121, InceptionV3, MobileNet, Resnet50, and Xception) are employed to classify autism images dataset into Binary classes. In this study, the size of all images in the dataset is  $224 \times 224 \times 3$ , The Adam optimizer is used to train each network architecture. Number of epochs is 35. The value of training and testing batch size is set to 32.

#### IV. Results

In this study, ASD diagnosis was presented as a binary classification issue by applying transfer learning technique to five pre-trained CNN models, the CNN models are (Densenet121, InceptionV3, MobileNet, ResNet-50, and Xception).all CNN architectures are completed after 35

epochs and batch size value of training and testing is 32. In this study the results showed that the maximum accuracy achieved 95.75% obtained by MobileNet model in 90.5 minutes and this is the shortest period of time of all models, and followed by Densenet121, InceptionV3, Xception, and ResNet-50. With 91.96%, 91.46%, 90.60%, and 72.41%. But the period of times of models are different and started with MobileNet model, InceptionV3, Densenet121, ResNet-50, and Xception. As shown in below Table 1.

Table .1 Comparison of pre-trained networks results.

Model	Accuracy	Time(minutes)
MobileNet	95.75%	90.5
Densenet121	91.96%	330.8
InceptionV3	91.46%,	239.7
Xception	90.60%	445.4
ResNet-50	72.41%.	346.1

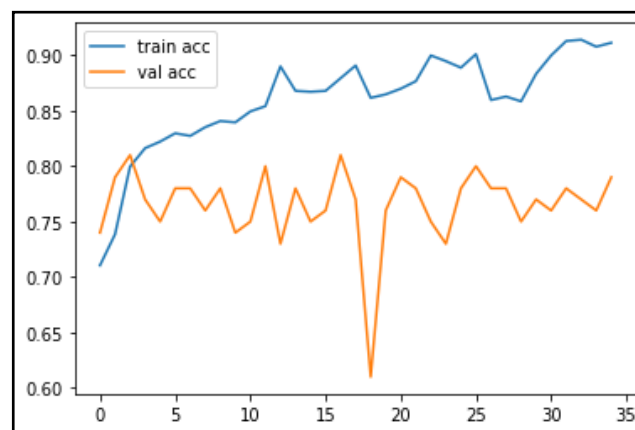


Figure 2: Densenet121 Training progress, Accuracy

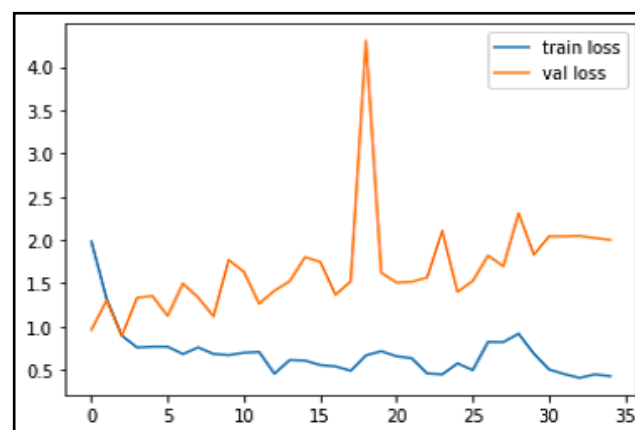


Figure 3: Densenet121 Training progress, Loss

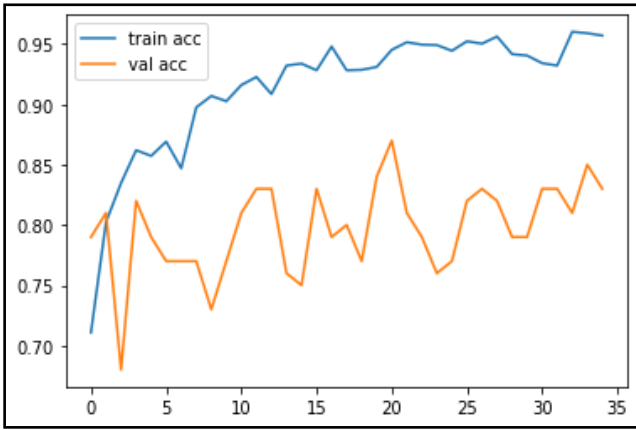


Figure 4: MobileNet Training progress, Accuracy

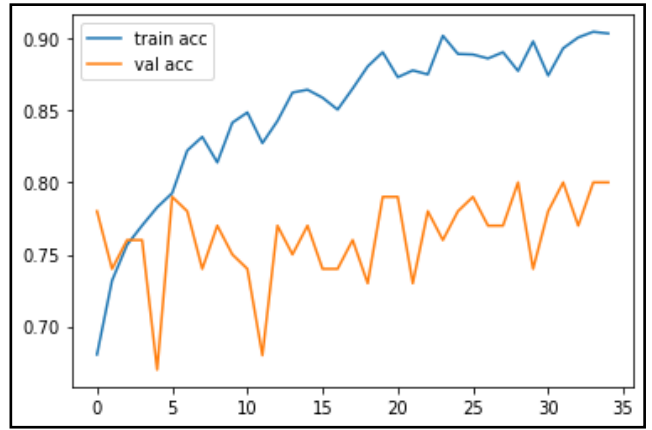


Figure 8: Xception Training progress, Accuracy



Figure 5: MobileNet Training progress, Loss

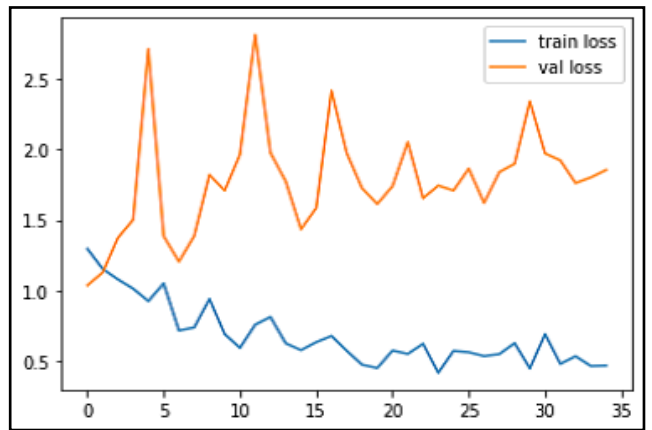


Figure 9: Xception Training progress, Loss

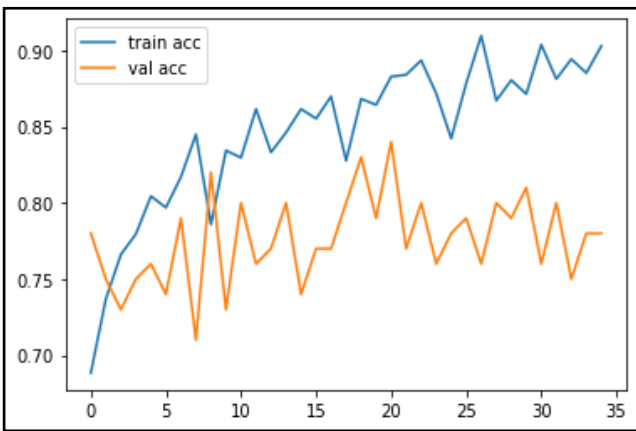


Figure 6: InceptionV3 Training progress, Accuracy

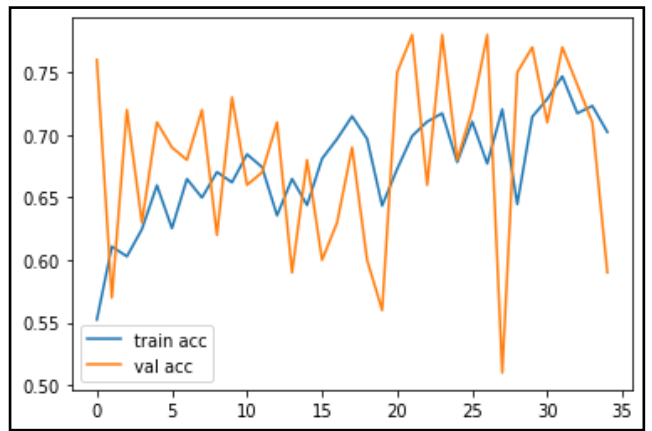


Figure 10: Resnet50 Training progress, Accuracy

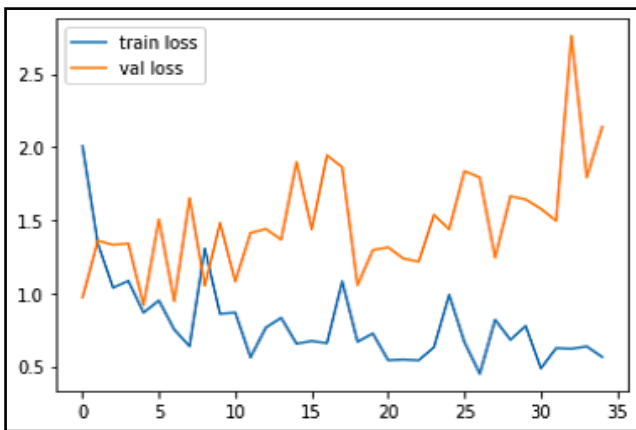


Figure 7: InceptionV3 Training progress, Loss

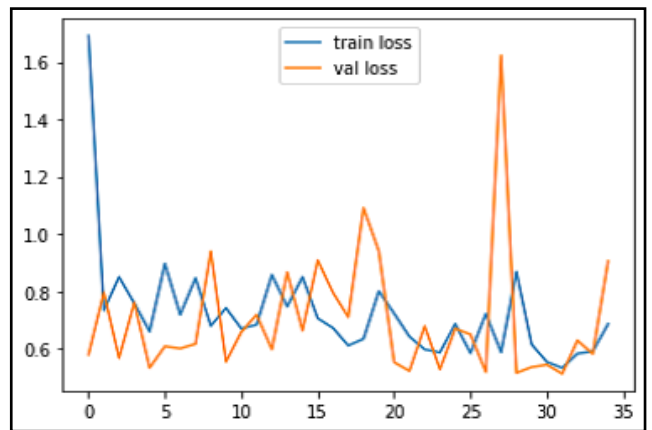


Figure 11: Resnet50 Training progress, Loss

## V. Conclusion

In this paper, we introduce a CNN based approach using transfer learning to detect and classify Autism Spectrum Disorder (ASD) patients from facial images as a binary classification issue. In this study five pre-trained CNN models are implemented with transfer learning technique on medical image dataset. The CNN models used to detect and classify the images are Densenet121, InceptionV3, MobileNet, ResNet-50, and Xception. All CNN models were trained used Adam optimizer and performed after 35 complete the value number of epochs. The batch size is 32. The results demonstrate that our model's accuracy with the test data is 95.75 %, which is best accuracy achieved on this dataset by the MobileNet model. In the future, we will compare the results of these models used in this study with some of the other models and refine the data with the proposed models. And trying more models on more medical images to increase the efficiency of classification to get best results.

## References

- [1] Nasser, I. M., Al-Shawwa, M., & Abu-Naser, S. S. (2019). Artificial Neural Network for Diagnose Autism Spectrum Disorder.
- [2] Sherkatghanad, Z., Akhondzadeh, M., Salari, S., Zomorodi-Moghadam, M., Abdar, M., Acharya, U. R., ... & Salari, V. (2020). Automated detection of autism spectrum disorder using a convolutional neural network. *Frontiers in neuroscience*, 13, 1325. Doland Drump, My Book, SYZ Publishers, Turkey, 2005.
- [3] Tamilarasi, F. C., & Shanmugam, J. (2020, June). Convolutional Neural Network based Autism Classification. In 2020 5th International Conference on Communication and Electronics Systems (ICCES) (pp. 1208-1212). IEEE.
- [4] Lai, Z., & Deng, H. (2018). Medical image classification based on deep features extracted by deep model and statistic feature fusion with multilayer perceptron. *Computational intelligence and neuroscience*, 2018.
- [5] Liu, W., Li, M., & Yi, L. (2016). Identifying children with autism spectrum disorder based on their face processing abnormality: A machine learning framework. *Autism Research*, 9(8), 888-898.
- [6] Erkan, U., & Thanh, D. N. (2019). Autism spectrum disorder detection with machine learning methods. *Current Psychiatry Research and Reviews Formerly: Current Psychiatry Reviews*, 15(4), 297-308.
- [7] Karabatak, M., Mustafa, T., & Hamaali, C. (2020, June). Remote Monitoring Real Time Air pollution-IoT (Cloud Based). In 2020 8th International Symposium on Digital Forensics and Security (ISDFS) (pp. 1-6). IEEE.
- [8] Beary, M., Hadsell, A., Messersmith, R., & Hosseini, M. P. (2020). Diagnosis of Autism in Children using Facial Analysis and Deep Learning. arXiv preprint arXiv:2008.02890.
- [9] Mythili, M. S., & Shanavas, A. M. (2014). A study on Autism spectrum disorders using classification techniques. *International Journal of Soft Computing and Engineering*, 4(5), 88-91.
- [10] Ali, N.A., Syafeeza, A.R., Jaafar, A., & Alif, M. (2020). Autism spectrum disorder classification on electroencephalogram signal using deep learning algorithm. *IAES International Journal of Artificial Intelligence*, 9(1), 91.
- [11] Rad, N. M., Kia, S. M., Zarbo, C., van Laarhoven, T., Jurman, G., Venuti, P., Marchiori, E., & Furlanello, C. (2018). Deep learning for automatic stereotypical motor movement detection using wearable sensors in autism spectrum disorders. *Signal Processing*, 144, 180-191.
- [12] Mohammed, S. H., & Çinar, A. (2021). Lung cancer classification with Convolutional Neural Network Architectures. *Qubahan Academic Journal*, 1(1), 33-39.
- [13] Mousa, A., Karabatak, M., & Mustafa, T. (2020, June). Database Security Threats and Challenges. In 2020 8th International Symposium on Digital Forensics and Security (ISDFS) (pp. 1-5). IEEE.
- [14] Karabatak, M., & Mustafa, T. (2018, March). Performance comparison of classifiers on reduced phishing website dataset. In 2018 6th International Symposium on Digital Forensic and Security (ISDFS) (pp. 1-5). IEEE.
- [15] Mustafa, T., & Varol, A. (2020, June). Review of the Internet of Things for Healthcare Monitoring. In 2020 8th International Symposium on Digital Forensics and Security (ISDFS) (pp. 1-6). IEEE.
- [16] Sharma, N., Jain, V., & Mishra, A. (2018). An analysis of convolutional neural networks for image classification. *Procedia computer science*, 132, 377-384.
- [17] Heinsfeld, A. S., Franco, A. R., Craddock, R. C., Buchweitz, A., & Meneguzzi, F. (2018). Identification of autism spectrum disorder using deep learning and the ABIDE dataset. *NeuroImage: Clinical*, 17, 16-23.
- [18] Hossain, M. D., Kabir, M. A., Anwar, A., & Islam, M. Z. (2021). Detecting autism spectrum disorder using machine learning techniques. *Health Information Science and Systems*, 9(1), 1-13.
- [19] Raj, S., & Masood, S. (2020). Analysis and Detection of Autism Spectrum Disorder Using Machine Learning Techniques. *Procedia Computer Science*, 167, 994-1004.
- [20] Shahamiri, S. R., & Thabtah, F. (2020). Autism AI: a New Autism Screening System Based on Artificial Intelligence. *Cognitive Computation*, 12(4), 766-777.

# Diabetic Retinopathy Classification from Retinal Images using Machine Learning Approaches

Indronil Bhattacharjee

Dept. of Computer Science and Engineering  
Khulna University of Engineering &  
Technology (KUET)  
Khulna, Bangladesh  
ibprince.2489@gmail.com

Al-Mahmud

Dept. of Computer Science and Engineering  
Khulna University of Engineering &  
Technology (KUET)  
Khulna, Bangladesh  
mahmud@cse.kuet.ac.bd

Tareq Mahmud

Dept. of Computer Science and Engineering  
Khulna University of Engineering &  
Technology (KUET)  
Khulna, Bangladesh  
hridaytareqmahmud@gmail.com

**Abstract**— Diabetic Retinopathy is one of the most familiar diseases and is a diabetes complication that affects eyes. Initially diabetic retinopathy may cause no symptoms or only mild vision problems. Eventually, it can cause blindness. So early detection of symptoms could help to avoid blindness. In this paper, we present some experiments on some features of Diabetic Retinopathy like properties of exudates, properties of blood vessels and properties of microaneurysm. Using the features, we can classify healthy, mild non-proliferative, moderate non-proliferative, severe non-proliferative and proliferative stage of DR. Support Vector Machine, Random Forest and Naive Bayes classifiers are used to classify the stages. Finally, Random Forest is found to be the best for higher accuracy, sensitivity and specificity of 76.5%, 77.2% and 93.3% respectively.

**Keywords**— Diabetic Retinopathy, Exudate, Blood Vessel, Microaneurysm, Random Forest.

## I. INTRODUCTION

People suffering from diabetes can have an eye complication called diabetic retinopathy. When blood sugar levels go high, that causes harm and erosion to the blood vessels in the retina. These affected blood vessels can fatten and exude. Alternately, the vessels may have been closed, may stop flowing bloods. Sometimes unnecessary and anomalous blood vessels starts to grow on the surface of the retina. These abnormal changes can damage one's vision, sometimes may destroy fully. According to severity of the disease, DR can be classified into two main stages. (a) Non-Proliferative Diabetic Retinopathy (NPDR) and (b) Proliferative Diabetic Retinopathy (PDR).

NPDR is the initial phase of diabetic retinopathy. Many individuals with diabetes suffers from it. With NPDR, tinier blood vessels excrete and fatten the retina. When the macula expands, this has been called macular edema. This is the most familiar reason why people having diabetes leads to blindness.

In case of NPDR, the blood vessels in the retina can clogged off too. This situation is named macular ischemia. When macular ischemia happens, macula cannot get the blood supply. Intermittently some minute particles called exudates can be grown in the retina. These affects one's vision too. If anybody suffers from NPDR, his eye sight will go blurry. Furthermore, NPDR is sub classed into 3 stages, Mild, Moderate and Severe.

Proliferate DR is the most critical phase of diabetic eye disease. It occurs when the retina starts growing excessive blood vessels, which is called neovascularization. These huge numbers of vulnerable vessels often bleed into the vitreous. When they only bleed a little, a few dark floaters are found. On the other hand, when they bleed a lot, that may block the whole of the vision.



Figure 1. Different stages of diabetic retinopathy (From top left): (a) Healthy Eye (b) Mild NPDR (c) Moderate NPDR (d) Severe NPDR (e) PDR

The objective of the paper are:

- Process color fundus retinal images for Diabetic Retinopathy detection.
- Extract key features from the pre-processed images.
- Detect the presence of Diabetic Retinopathy.
- Classify whether the Diabetic Retinopathy is Proliferative or Non-proliferative.

## II. THE PROPOSED SYSTEM

**Input:** Colour fundus retinal images

**Output:** Diabetic Retinopathy is not present, Mild, Moderate, Severe or PDR

**Process:**

- Step 1: Input the initial fundus image
- Step 2: Preprocess the initial image
- Step 3: Optical disk removal
- Step 4: Exudates detection
- Step 5: Blood Vessels detection
- Step 6: Microaneurysm detection

- Step 7: Features extraction
- Step 8: Apply to classifiers
- Step 9: Classify the Diabetic Retinopathy stages
- Step 10: Detect whether it is Healthy, Mild, Moderate, Severe or PDR eye

### III. METHODOLOGY

#### A. Dataset

To evaluate our method, we have used a dataset named as Diabetic Retinopathy (Resized) from Kaggle. The dataset has a total of 13402 retinal images and corresponding levels of Diabetic Retinopathy for each image.

#### B. Data Preprocessing

Data preprocessing has been done in two steps, general preprocessing for all the images and specific preprocessing for individual feature extraction.

##### 1) General Preprocessing:

- *Resizing:* In this work, the sizes of the actual images in the dataset were 1024x1024 pixels. As the dataset is huge in size, the images have been stored with size 350x350 pixels for reducing the computational time.

- *Green Channel Extraction:* Preprocessing has been done with the aim of improving the contrast level of the fundus images. For contrast enhancement of the retinal images, some components like the red and blue components of the image were commonly discarded before processing. Green channel shows the clearest background contrast and greatest contrast difference between the optic disc and retinal tissue. Red channel is comparatively lighter and vascular structures are visible. The retinal vessels are lightly visible but those show less contrast than that of the green channel. Blue channel contains very little information and is comparatively noisier.

- *Contrast Limited Adaptive Histogram Equalization:* Contrast Limited Adaptive histogram equalization (CLAHE) is used for enhancing the contrast level of the images. CLAHE calculates different histograms of the image and uses these information to reallocate intensity value of image. Hence, CLAHE is significant for improving the regional contrast and enhancement of the edges in all the regions of an image.

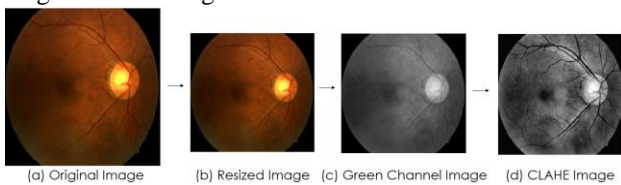


Figure II. General preprocessing

##### 2) Specific Preprocessing:

- *Exudate Detection:* Firstly, optical disc has been removed using red channel of the image. Then, using 6x6 ellipse shaped structuring element, morphological dilation is applied. Non linear median filter is used for noise removal. Exudates are in high intensity values. So it

has been extracted using thresholding. After applying these preprocessing, pixels having intensity value higher than 235 are set to 255 and the rest of them are set to 0 for having the clearest view. Then traversing the image, area of exudates are calculated. The images of different steps are illustrated in Figure III.

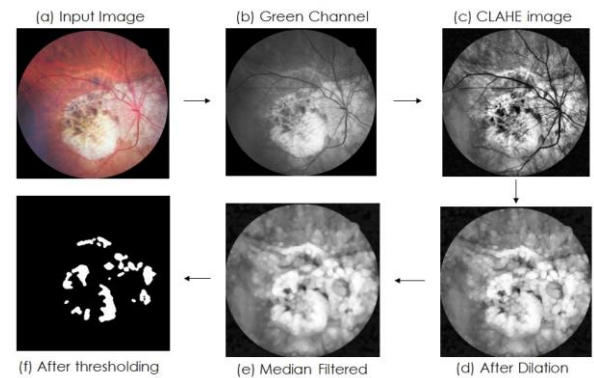


Figure III. Preprocessing for Exudate detection

- *Blood Vessel Extraction:* Blood vessel is one of the most important features for differentiating diabetic retinopathy stages. After obtaining the green channel image and improving the contrast of the image, several steps has been done for extracting blood vessel. Alternate sequential filtering (three times opening and closing) using three different sized and ellipse shaped structural element 5x5, 11x11 and 23x23 is applied on the image. Then the resultant image is subtracted from the input image. Subtracted image has lots of small noises. Those noises are removed through area parameter noise removal. Contours of each components including noises are found by finding out the contours and calculating the contour area and remove the noises which are comparatively bigger in size (200 used as reference). Then the resultant image is binarized using a threshold value. Finally the number of pixels that covers the blood vessels area are calculated. The images of different steps are illustrated in Figure IV.

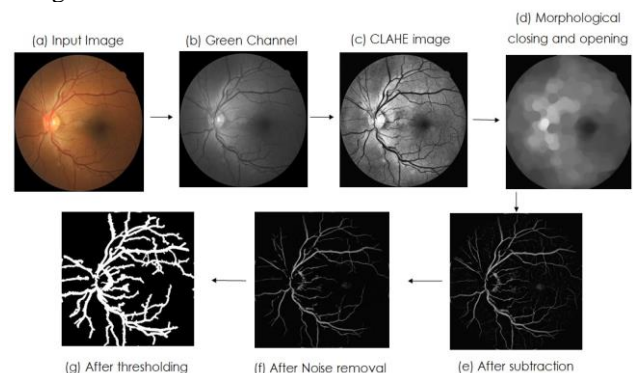


Figure IV. Preprocessing for Blood Vessel detection

- *Microaneurysm Extraction:* Green component is applied to extract microaneurysm. For better contrast, CLAHE is used. Then median filter is used for noise

removal. 7x7 ellipse shaped structural element is used for morphological operation. Morphological operation erosion is applied and then the image is inverted. For joining the disjoint segments of blood vessel, morphological closing is used. Then the image has been binarized. As the blood vessel, haemorrhage and microaneurysm is having the almost same intensity, all these components will be detected altogether in the binarized image. Since microaneurysm is smaller in size, it has been extracted using contour area. The images of different steps are illustrated in Figure V.

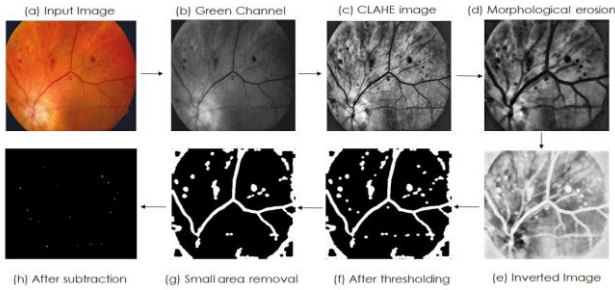


Figure V. Preprocessing for Microaneurysm detection

### C. Dataset Splitting

The dataset has been divided into two parts, where 75% as training data and 25% as test data. Therefore, 10052 training images has been used to train the model and it has been tested on 3350 images.

### D. Data Scaling

In this system, standard scalar has been used to scale all the data to limit the ranges of the variables. Using data scaling, those can be compared on the exact environments in case of all the algorithms.

### E. Selection of Features

Feature extraction will be done from preprocessed images shown in Figure III, IV and V. The features which are extracted to detect Diabetic Retinopathy are-

- Histogram of Exudates
- Zeroth Hu moment of Exudates
- Histogram of Blood Vessels
- Zeroth Hu moment of Blood Vessels
- Histogram of Microaneurysm
- Zeroth Hu moment of Microaneurysm

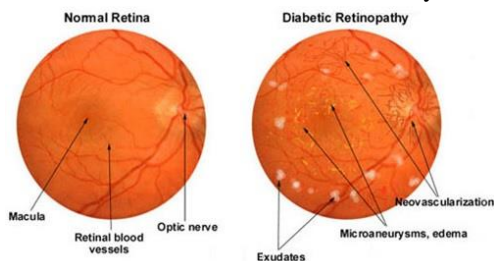


Figure VI. Differences between a normal retina and DR affected retina in terms of different features

### F. Classification

Prediction has been performed using Support Vector Machine (SVM), Random Forest (RF) and Naive Bayes classifiers.

#### 1) Random Forest:

Random Forest (RF) is an ensemble tree-based learning algorithm. The RF Classifier could be a set of decision trees from an arbitrarily chosen subset of the preparing sets. It totals the votes from distinctive choice trees to choose the ultimate class of the test question. The elemental concept behind RF classifier could be a basic but capable one - the intelligence of swarms. An expansive number of moderately uncorrelated trees working as a committee will beat any of the person constituent models. Uncorrelated models can deliver gathering expectations that are more precise than any of the individual predictions. The reason for this wonderful impact is that the trees ensure each other from their personal mistakes as long as they don't always all mistakes within the same heading. Whereas a few trees may be off-base, numerous other trees will be right, so as a gather the trees are able to move within the adjusted heading.

#### 2) Support Vector Machine:

SVM classify the input images into two classes such as Diabetic Retinopathy affected eye and normal eye using its features. As SVM is a binary classifier, our first task is to classify which eye is affected by Diabetic Retinopathy and which is a healthy one. After first classification, our next task is to use Support Vector Machine again. This time it is applied only on the affected ones. It will again classify which Diabetic Retinopathy is non-proliferative i.e. is in initial stage and which on is in proliferative i.e. is in severe state.

Support Vector Machine has been utilized since the SVM is based on a convex objective function that never stuck into the local maxima. The ideal hyperplane is the shape of the isolating hyperplane and the objective work of the optimization issue does not depend unequivocally on the dimensionality of the input vector but depends as it were on the inward items of two vectors. This fact permits developing the isolating hyperplanes in high-dimensional spaces.

#### 3) Naïve Bayes:

Naive Bayes classifier isn't a single algorithm, but a collection of algorithms where all of them contains a common rule, that is, each match of highlights being classified is free of each other. Naive Bayes is mainly an ensemble algorithm based on Bayes' Theorem.

### G. Evaluation Metrics

Accuracy, Sensitivity and Specificity are used as evaluation metrics of the model. Accuracy, Sensitivity and Specificity are calculated using (1), (2) and (3) respectively.

$$Accuracy = \frac{True\ Positive + True\ Negative}{Total\ Number\ of\ Data} \quad (1)$$

$$Sensitivity = \frac{True\ Positive}{True\ Positive + False\ Negative} \quad (2)$$

$$Specificity = \frac{True\ Negative}{True\ Negative + False\ Positive} \quad (3)$$

Finally, the workflow of this paper is illustrated in Figure VII

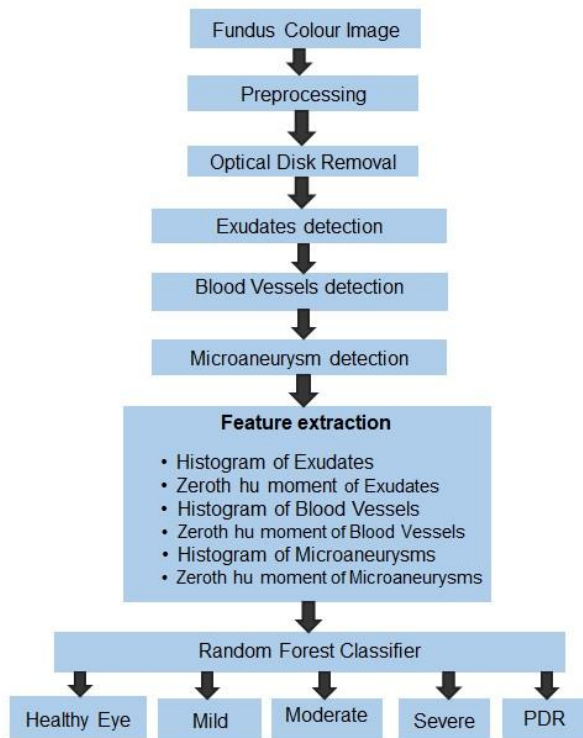


Figure VII. Flow Diagram of the system

#### IV. RESULTS

Comparison of evaluation metrics for Random Forest, Support Vector Machine and Naïve Bayes classifier is shown in Table 1.

TABLE I. PERFORMANCE MEASURES OF DIFFERENT CLASSIFIERS

Classifier	Accuracy (%)	Sensitivity (%)	Specificity (%)
Random Forest	76.5	77.2	93.3
SVM	70.2	72.5	82.0
Naïve Bayes	67.3	69.4	75.2

In the experiment, Random Forest gives the highest accuracy, sensitivity and specificity. That is why Random Forest has been chosen as the best of the three classifiers used in this work.

Average sensitivity of the classification = 77.2%  
 Average specificity of the classification = 93.3%  
 Accuracy of the classification = 76.5%

Predicted	Actual				
	Normal	Mild	Moderate	Severe	PDR
Normal	989	100	114	63	66
Mild	116	466	10	6	6
Moderate	122	10	492	8	10
Severe	82	2	12	316	8
PDR	40	4	14	4	291

Figure VIII. Confusion Matrix for the five classes

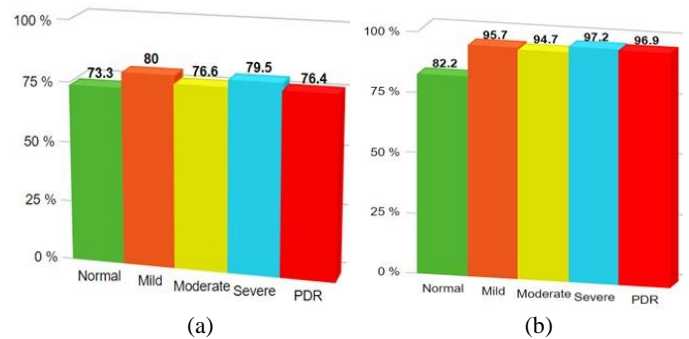


Figure IX. Bar diagram of (a) sensitivity and (b) specificity for each class

The evaluation measures are compared with the related works in terms of number of classes, methods used, features etc.

TABLE II. COMPARISON OF DIABETIC RETINOPATHY DETECTION BY VARIOUS RESEARCHERS

Reference	Number of Classes	Method	Accuracy (%)	Sensitivity (%)	Specificity (%)
Sinthanayothin et al. [1]	2	Moat operator	Not reported	70.2	70.6
Singalavanija et al. [2]	2	Exudate, Blood vessel, Microaneurysm	Not reported	74.8	82.7
Kahai et al. [3]	2	Decision support system	Not reported	80	63
Wang et al. [4]	4	Area of blood vessel	74	81.7	92
Nayak et al. [5]	3	Blood vessel, exudates, texture	73.6	70.3	90

Acharya et al. [6]	5	Higher- order spectra	72	82.5	88.9
Lim et al. [7]	5	Blood vessels, exudates, microaneurysm, haemorrhage	75.9	80	86
This work	5	Histograms and Zeroth Hu moments of blood vessel, exudates, microaneurysm	76.5	77.2	93.3

Sinthanayothin et al. [1] distinguished Diabetic Retinopathy using image processing techniques from a healthy retina. In this proposed system, fundus images were preprocessed using adaptive local contrast enhancement. This method, established on a multilayer neural network, made 70.21 percent sensitivity and 70.66 percent specificity.

Kahai et al. [3] developed a system for the initial identification of the Diabetic Retinopathy. The identification system is based on a testing problem of binary-hypothesis that results only yes or no. Bayes optimization criterion was used to the raw fundus images for the initial identification of the DR. This method was able to detect the appearance of microaneurysms having sensitivity of 80 percent and specificity of 63 percent correctly.

Wang et al. [4] have classified healthy, moderate, and severe DR stages using morphological image processing approaches and a feedforward deep learning network. In this system, the existence and covering region of the components of the lesions and blood vessels are selected as the main features. The classification efficiency of this work was 74 percent, the sensitivity was 81 percent, and the specificity was 92 percent. Featuring the lesions and blood vessels, and texture parameters, they classified the input images into healthy, Moderate, and Severe DR [4].

An automatic identification system of DR was proposed by Acharya et al. [6]. They classified normal, mild, moderate, severe and PDR using the two spectral layered invariant features of higher-order spectra approaches and a Support Vector Machine classifier [8]. This work reported an accuracy of 72%, a sensitivity of 82%, and a specificity of 88%.

In this work, the fundus images are classified into five classes using the histograms and the zeroth Hu moments of the exudates, microaneurysms, and blood vessels present in the eye. Random Forest is used for the classifier. The classifier is able to identify the unknown class accurately with an efficiency of more than 76.5 percent with sensitivity 77.2 percent and specificity 93.3 percent.

Since, the instances of each class are important in this work, we have calculated the evaluation metrics by using macro-averaging method. Because macro average reveals the better scenario of the smaller classes and it is to the point and more accurate when performances on each and every classes are important equally.

## V. CONCLUSION

After studying the existing systems, we conclude that our proposed technique is successfully detecting Diabetic Retinopathy. Along with this, the proposed method is classifying into five classes of Diabetic Retinopathy. Classification has been done based on three features- area of exudates, area of blood vessel and area of microaneurysm. And using this features, we have classified into five classes as normal eye, mild NPDR, moderate NPDR, severe NPDR and PDR. Using Random Forest classifier, we have gained accuracy= 76.5%, sensitivity= 77.2% and specificity= 93.3%. The metrics we have found in this work are compared with the existing works.

In this paper, we have performed the Diabetic Retinopathy classification using Random Forest classifier with some essential features like exudates, blood vessel and microaneurysm. In future, we hope to make it work for some more classifiers like K-Nearest Neighbor classifiers and so on using some secondary features like haemorrhage also. Moreover, we can perform this classification method using larger dataset of infected eyes using neural network model in future.

## REFERENCES

- [1] C. Sinthanayothin, J. F. Boyce, T. H. Williamson, H. L. Cook, E. Mensah, S. Lal, and D. Usher, "Automated detection of diabetic retinopathy on digital fundus images" *Diabetic Medicine*, vol. 19, no. 2, pp. 105–112, 2002.
- [2] Singalavanija, A., Supokavej, J., Bamroongsuk, P., Sinthanayothin, C., Phoojaruenchanachai, S., and Kongbunkiat, V. Feasibility study on computer-aided screening for diabetic retinopathy. *Jap. J. Ophthalmology*, 2006, 50(4), 361–366.
- [3] P. Kahai, K. R. Namuduri, and H. Thompson, "A decision support framework for automated screening of diabetic retinopathy." *Hindawi Publishing Corporation*, Feb 02, 2006.
- [4] H. Wang, W. Hsu, K. Goh, and M. Lee, "An effective approach to detect lesions in color retinal images," vol. 2, 02 2000, pp. 181 – 186 vol. 2.
- [5] Nayak, J., Bhat, P. S., Acharya, U. R., Lim, C. M., and Kagathi, M. Automated identification of different stages of diabetic retinopathy using digital fundus images.
- [6] U. R. Acharya, E. Y. K. Ng, J. H. Tan, V. S. Subbhuraam, and N. Kh, "An integrated index for the identification of diabetic retinopathy stages using texture parameters," *Journal of medical systems*, vol. 36, pp. 2011–20, 02 2011.
- [7] C M Lim, U R Acharya, E Y K Ng, C Chee and T Tamura, *Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine*, 2009 223: 545
- [8] Osareh, A.; Mirmehdi, M.; Thomas, B.T.; Markham, R. In *MICCAI'02: Proceedings of the 5th International Conference on Medical Image Computing and Computer-Assisted Intervention-Part II*; Springer-Verlag: London, UK, 2002; pp 413–420



# Methodology for Syngas Energy Assessment

Melda Ozdinc Carpinlioglu  
Gaziantep University  
Department of Mechanical Engineering  
Gaziantep, Turkey  
melda@gantep.edu.tr

**Abstract**— Energy assessment of syngas produced by solid granulated fuel decomposition in a special microwave plasma gasification process is the topic of presentation . Instantaneous measurements of syngas volumetric content covering the total gasification time of the process and gas chromatography analysis on the collected sample syngas amount during approximately 30% of total gasification time are compared . Instantaneous volumetric content of syngas and the gasification time and location averaged process temperature which is defined as syngas temperature ,  $T_{syn}$  are the major energy assessment parameters. Gas chromatography analysis verifies the transient nature of syngas production..

**Keywords**— syngas, volumetric content, gas chromatography, transient, instantaneous measurement

## I. Introduction

The criticism on the literature and state of art under the light of a research between 2015-2018 regarding microwave plasma gasification in a test system MCw GASIFIER for biomass conversion to syngas are available in [1,2,3,4,5]. A review on the manner [1] with a methodology on thermodynamical treatment of the process [2] can be treated as the pre-research publications. The article [3] is on an overall treatment of research discussing syngas generation with a variety of fuel as a post-research article. A Ph. D study [4] was completed with the support of the research [5]. In terms of the current efforts in the relevant field a proceedings paper was presented in May 2021[6]. It is denoted [6] that research on microwave MCw plasma gasification is onwards for recycling of biomass and waste to energy conversion due to the importance of the manner in terms of the advantages of MCw plasma gasification and existing gaps in physical nature of the process .[7,8,9,10,11] The major attention in this paper is given to the methodology for the energy content determination of syngas produced by MCw plasma gasification of solid granulated fuel. The operational results of MCw GASIFIER are referred for the purpose. The available data on instantaneous volumetric content of generated syngas and temperature measurements during the total decomposition period are used with the data on gas chromatography analysis of stored syngas samples before the termination of the decomposition.

The discussion is based on a brief on the operational test cases of MCw GASIFIER [1-5]. Syngas energy assessment is described with the critical parameters . It is also aimed to determine the functional relationships between selected parameters expressing the physical process of gasification.

## II. Brief on process, basic parameters and methodology

The operation of MCw GASIFIER[5] was such that air at standard temperature and pressure condition was used as plasma carrier. Granulated particles of different types of biomass having the particle size range of 0.1 mm – 1 mm were loaded as a static fuel bed in the reactor. Coal, C

sawdust, SD and polyethylene pellets, PP were used fuel. Syngas was defined as the gasified amount of fuel excluding ash left the process but including the supplied amount of air. Therefore syngas production is governed by the nature of fuel and MCw plasma decomposition process (Figure I)

Monitoring of gasification was based upon the instantaneous measurements , IM of local instantaneous temperatures,  $T$  along the reactor and syngas volumetric content . The local instantaneous temperature measurements were taken by B type (Pt18 Rh- Pt) thermocouples located along the reactor at specified 5 locations. Thermocouples had a sensitivity of  $\pm 4$  C. Local temperature measurement was at 1 second sampling frequency . A semi-continuous commercial gas analyzer called MRU-VARIO PLUS was used to measure the syngas volumetric content . CO, CO<sub>2</sub>, CH<sub>4</sub>, H<sub>2</sub>, N<sub>2</sub> and O<sub>2</sub> in syngas was determined with an accuracy of up to 100 % and up to 25 % respectively. The syngas volumetric content measurement was at 2 seconds sampling frequency . Fuel decomposition was through a defined process –gasification time ,  $t_g$  . Instantaneous measurements were executed from  $t = 0$  to  $t = t_g$  . The gasification time ,  $t_g$  was resulted in no solid fuel left in the reactor with a minute amount of ash filtered from generated syngas through a cyclone separator located in front of the MRU-VARIO PLUS gas analyzer.

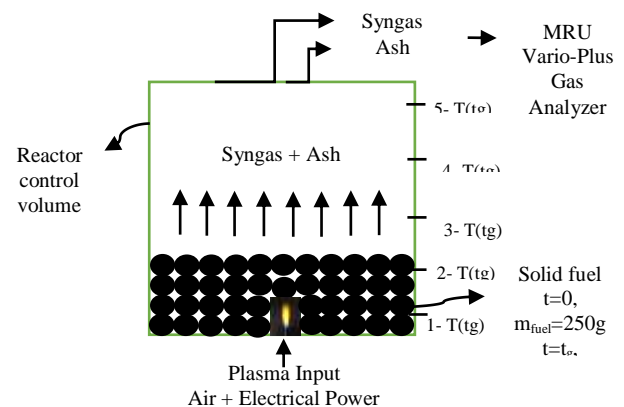


Figure I Sketch of the process [5]

Stored amount of syngas during fuel decomposition before the termination of the process was collected in bags during the so called storage time,  $t_s$ . The chromatographic analysis , GC of stored syngas samples was done in TUBITAK-MAM Laboratory according to TS EN ISO 6974 and TS EN ISO 6975 standards . Atmospheric pressure and reference standard temperature  $T_R = 20^\circ C$  were valid in the analysis. C1-C6 hydrocarbons( Ethane, Ethylene, Propylene ,I-Pentane n- Hexane ,Butene , etc ) besides CO, CO<sub>2</sub>, CH<sub>4</sub>, H<sub>2</sub>, N<sub>2</sub> and O<sub>2</sub> components were determined .

The local instantaneous temperatures during fuel decomposition process,  $t_g$  were varying as a function of fuel and process. Similarly volumetric amounts of CO, CO<sub>2</sub>, CH<sub>4</sub>, H<sub>2</sub>, N<sub>2</sub>, O<sub>2</sub> in syngas were also varying during  $t_g$ . Therefore process-decomposition –reactor temperature can be defined as syngas temperature,  $T_{syn}$  as an ensemble averaged parameter using collected local instantaneous temperature data during  $t_g$ . Similarly molecular weight of syngas  $M_{syn}$  as a secondary ensemble averaged parameter can be calculated using collected syngas volumetric content data during  $t_g$ . The energy content of syngas can be given by HHV of syngas. HHV<sub>syn</sub> can be calculated based upon Dalton-Amagat Model on instantaneous –ensemble averaged syngas volumetric content similar to the calculation of  $M_{syn}$ . The calculations are based upon the assumption of perfect gas treatment for syngas and referring to  $T_{syn}$  [1,2,3,5]. HHV<sub>syn</sub>, relative density, RD and Wobbe Index (WI) of syngas [12,13] can be determined in GC analysis of stored syngas sample. GC analysis is based upon the treatment of perfect gas and real gas assumptions for syngas separately. The determination of gross and net quantities of the parameters are also available [5]. Table I lists the defined parameters. HHV<sub>syn</sub> is the common parameter in different methodology.

Table I Parameters for syngas energy assessment derived from IM ,GC analysis

Parameter	Base	Definition	Explanation
$t_g$	IM	Process duration	No solid fuel left in reactor
$t_s$	GC	Storage time	$t_s < t_g$
$T_{syn}$	IM	$t_g$ based Process-Reactor –Syngas temperature	Ensembled averaged local temperatures in reactor during $t_g$
$M_{syn}$	IM	Molecular Weight of Syngas	Calculated Dalton-Amagat Model
HHV <sub>syn</sub>	IM	Higher Heating Value of Syngas	Calculated Perfect Gas Mixtures at $T_{syn}$
HHV <sub>syn</sub>	GC	Higher Heating Value of Syngas	Calculated Perfect Gas Real Gas
RD <sub>syn</sub>	GC	Relative density of syngas	Syngas density / Air density Perfect Gas Real Gas
WI <sub>syn</sub>	GC	Wobbe Index	Amount of Energy from syngas conversion $HHV / RD^{0.5}$ Perfect Gas Real Gas

### III. Results and Discussion

The produced syngas for the gasification of polyethylene pellets, PP as solid fuel with 4.8 kW microwave power use at 50 standard L/min of air use are referred. IM data

gathered during the total duration of process,  $t_g$  and GC analysis data conducted on a stored sample syngas amount during  $t_s < t_g$  with  $t_s/t_g = 0.3$  corresponding to a partial gasification are compared. The comparison is given in terms of syngas volumetric content in part A.

Furthermore syngas energy assessment for the gasification of polyethylene pellets, PP as solid fuel with a variable microwave power 3 kW, 4.8 kW and 6 kW use at 50 standard L/min of air are referred using data gathered by IM and GC analysis separately and in a combined manner in part B.

Influence of fuel content on the generated syngas is presented in terms of proposed routes for fuel decomposition in part C.

#### A. Syngas Volumetric Content Data via GC and IM

GC analysis data are such that Ethane, Ethylene, Propylene, I-Pentane, n-Hexane, Butene, Acetylene, 1-3 Butadiene, Propane, n-Hexane etc. in syngas can be determined. However the amounts of these components are rather small referring to the maximum 2.887% of Ethylene and minimum 0.002% of n-Hexane. In both methods the determined common components of syngas are CH<sub>4</sub>, O<sub>2</sub>, N<sub>2</sub>, CO<sub>2</sub>, CO, H<sub>2</sub>. The comparison of the common components in syngas for the referred data sets can be discussed as follows:

1. CH<sub>4</sub> content in GC data (1.94%) is in the same order of IM data (1.26%).

2. O<sub>2</sub> and N<sub>2</sub> content in GC data are such that O<sub>2</sub> (4.132%) and N<sub>2</sub> (75.653%) while in IM data O<sub>2</sub> (1.44%) and N<sub>2</sub> (51.5%) are determined.

3. CO<sub>2</sub>, CO, H<sub>2</sub> contents of IM data CO<sub>2</sub> (13.01%), CO (21.31%), H<sub>2</sub> (10.80%) are greater than the corresponding magnitudes of GC data as given with CO<sub>2</sub> (6.55%), CO (4.55%), H<sub>2</sub> (1.91%). The instantaneously determined magnitudes of CO<sub>2</sub>, CO, and H<sub>2</sub> have approximately 2, 4.3 and 5.7 multiples of the ones determined by GC.

The reason of the observed deviations is due to the incomplete decomposition of fuel with  $t_s < t_g$ . The similar magnitudes of CH<sub>4</sub> of syngas determined by IM and GC data can be due to the completed CH<sub>4</sub> formation for the time period  $t = t_s$ . The greatest magnitudes of O<sub>2</sub> and N<sub>2</sub> by GC data also confirm that the gasification procedure is not completed for  $t_s < t_g$  since O<sub>2</sub> and N<sub>2</sub> are the components of air supplied steadily and continuously while fuel decomposition sensed, in terms of formation of CO<sub>2</sub>, CO, CH<sub>4</sub>, H<sub>2</sub> of syngas is not completed.

#### B. Syngas Energy Assessment :Comparison via GC and IM

Syngas energy level is a function of generated syngas amount, its volumetric content and temperature,  $T_{syn}$ . The parameters describing the energy level are given in Table I. The changes in the syngas content described in part A are due to the measurement and analysis conditions. In reference to the discussed cases for the gasification of polyethylene pellets  $t_s/t_g$  is between 0.28-0.35 (Table II). This means that collected sample mass for GC is the one before the complete decomposition of the fuel. The process of decomposition of fuel is not a steady state steady flow process. Therefore IM at varying high temperatures are the

correct ones resembling the physical nature of the process. Furthermore the temperature acting in IM is  $T_{syn}$ . It is under the effect of decomposition since ensembled average temperature during process at different locations is used. GC analysis is based on reference temperature,  $T_R$  and  $T_{syn}$  is  $\gg T_R$ . Therefore the difference of  $T_{syn} - T_R$  is a dominant term in discussion.

Table II Syngas Energy based upon GC data for polyethylene pellets ,PP gasification at different power use in comparison with IM data

$t_s / t_g$	GC RD	GC Net WI MJ/m <sup>3</sup>	GC E <sub>syn</sub> (kJ) From WI
0.28	1.05	8.21	7563
0.3	0.98	5.41	4395
0.35	0.97	7.44	5582
	IM WI MJ/m <sup>3</sup>	IM E <sub>syn</sub> (kJ)	IM $M_n = M_{syn} / M_{air}$
0.28	2.89	3298	0.96
0.3	2.713	3274	0.94
0.35	2.573	3327	0.91

Net WI value and RD value of generated syngas produced with gasification of PP at different power use based on GC analysis are not varying depending on perfect gas or real gas treatment of syngas. (Table II) The maximum magnitude of WI is belonging to 3 kW power use with 8.21 MJ/m<sup>3</sup> with the values 5.41 MJ/m<sup>3</sup> and 7.44 MJ/m<sup>3</sup> corresponding to 4.8 kW and 6 kW power use respectively. RD<sub>syn</sub> is in the order of 1 meaning density of syngas similar to the one of standard air. Instantaneous measurements are used to determine molecular weight of syngas. Calculated molecular weight of syngas,  $M_{syn}$  is given as a ratio of atmospheric standard air molecular weight,  $M_{air}$  in Table II as normalized molecular weight,  $M_n = M_{syn} / M_{air}$ . The magnitudes of  $M_n$  are in the same order of RD. This means that almost independent of method used, the produced syngas has similar of air in terms of density and molecular weight. Presence of air in syngas may be the reason. Therefore the respective orders of O<sub>2</sub> and N<sub>2</sub> in syngas can be determined. The ratio of O<sub>2</sub>/N<sub>2</sub> of syngas,  $O_2/N_2_{syn}$  with that of air is defined as normalized ratio of O<sub>2</sub>/N<sub>2</sub>,  $(O_2/N_2)_n$ . The variation of  $(O_2/N_2)_n$  with  $O_2/N_2_{syn}$  is shown in Figure II using all of the collected data with C, SD and PP gasification. The syngas generation by all kinds of fuel gasification almost follow the single line shown. The data behaviour in Figure II means that oxygen in syngas is mostly less than its amount in standard pure atmospheric air. As can be seen only 2 data with C gasification  $(O_2/N_2)_n > 1$ . The presence of CO, CO<sub>2</sub> components in syngas causing reduction in pure O<sub>2</sub> is the possible reason of the fact.

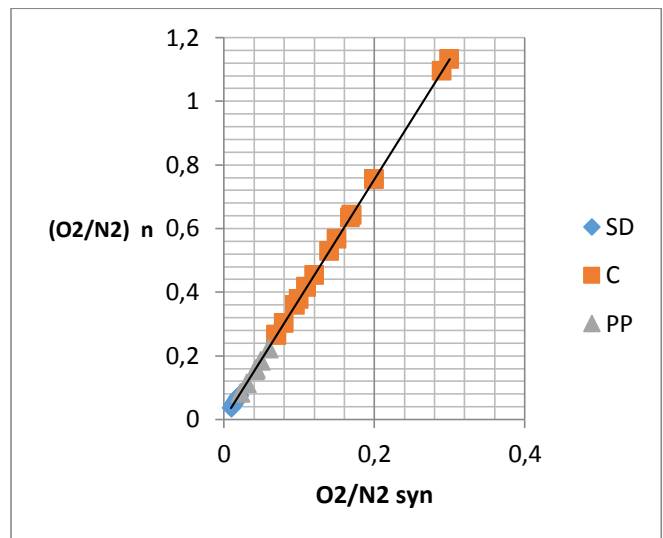


Figure II O<sub>2</sub>/N<sub>2</sub> of generated syngas expressed relative to air

As a further comment  $M_n$  is given with  $M_{syn}$  in Figure III in reference to all of the collected data with C, SD and PP gasification. The magnitudes of  $M_n$  are between 0.9 and 1.15. All data roughly independent of fuel type following a single line shown.

Table II lists syngas energy,  $E_{syn}$  calculated based upon WI, RD of GC data using the amount of syngas measured from the operational cases of IM. The assumption is based upon the similarity of syngas generation with decomposition of fuel for  $t > t_s$  since GC data is referred in calculation. However the calculated values with IM data at different power use are in the same range of  $E_{syn} = 3200$  kJ. The values of  $E_{syn}$  derived from GC are dependent on power use and varying in the approximate order of 4000-7300 kJ. In fact combined data use from GC and IM is one of the reason. The second reason is the temperature difference,  $T_{syn} - T_R$  which induces sensible energy increase for syngas in GC analysis. The magnitudes of  $E_{syn}$  determined by WI utilization are greater than the magnitudes obtained by IM.  $E_{syn}$  calculated by IM are 0.43, 0.74 and 0.59 of  $E_{syn}$  of GC.

WI defined as a GC analysis parameter is also calculated using data gathered by IM as an alternative method. The results are given in Table II with the approximate magnitudes of 2.89, 2.7, 2.57 MJ/m<sup>3</sup>. WI magnitudes derived from IM are almost 0.35, 0.5 and 0.34 of the corresponding magnitudes of GC analysis.

Therefore different methods give different values for  $E_{syn}$ . The combined method utilization seems to be not realistic due to the difference between  $t_s$  and  $t_g$ , temperature difference of  $T_{syn} - T_R$  and the difference in the amount of syngas.

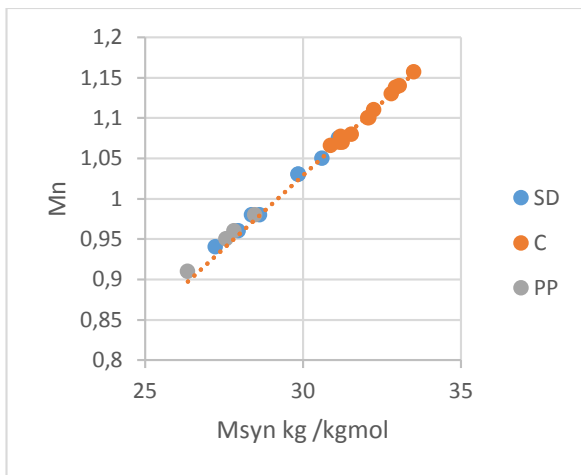


Figure III Variation of  $M_n$  with  $M_{syn}$

### C. Influence of Fuel Content on Syngas Content: Analysis of Decomposition

In order to describe the influence of fuel on the generated syngas composition and the thermochemical decomposition process Figure IV and Figure V can be referred for a final deduction just as samples. IM data are used.

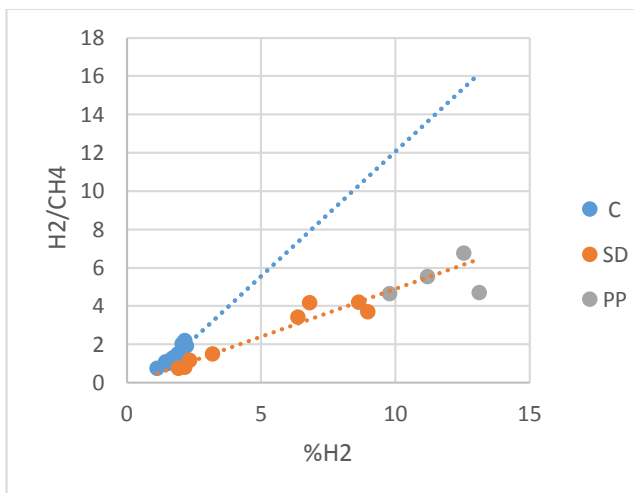


Figure IV Influence of fuel on the generated syngas  $H_2/CH_4$  variation with %  $H_2$  content

In these figures the ratio of  $H_2/CH_4$  in syngas versus %  $H_2$  in syngas and the ratio of  $CO/CO_2$  versus %  $CO_2$  in syngas are used. The selection of the ratios comes from the formation of the relevant components are the measure of fuel type with varying C, H, O amounts and fuel based thermochemical decomposition process. As a common reference base in Figure IV and Figure V %  $H_2$  and %  $CO_2$  in syngas are taken as dependent components. All of the collected data with C, SD and PP gasification are used.

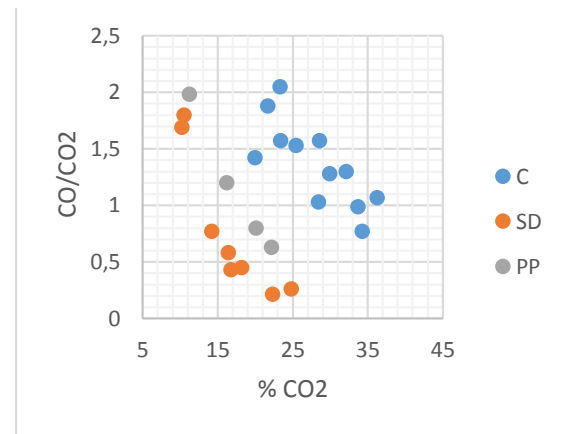


Figure V Influence of fuel on the generated syngas  $CO/CO_2$  with %  $CO_2$  content

It is seen that increase in  $H_2\%$  content of syngas is coupled with an increase in  $H_2/CH_4$  ratio of syngas. Meanwhile increase in  $CO_2\%$  of syngas is coupled with a decrease in  $CO/CO_2$  ratio of syngas. The increase in dependent parameter given in the abscissa and the used ratio definitions is the logical fact for the observed data behaviour. Therefore Figures IV and V can be interpreted as the sample plots on thermochemical decomposition of fuel which is a severe function of fuel type.

### Acknowledgments

The author expresses her gratitude to TUBITAK for the completed research project with a grant number of 115M389 (2015-2018). TUBITAK –MAM Gebze Laboratory personnel for the conducted gas chromatography analyses of the samples and the criticism of the referees of the 115 M 389 project during 2015-2018 period also deserve thanks. The assistance of Dr. A.Sanlisoy in the project duration is acknowledged. The hardworking of Res. Assistant E.Ozkur in revision process is welcome. The author lastly but not leastly thanks for the criticism and comments of the referees.

### III. Conclusion

It seems that syngas energy should be expressed through instantaneous gasification based measurements IM and gas chromatography GC analysis. GC analysis on the sample gas from gasification process is confirming the transient nature of the process giving the total available energy in generated syngas due to the difference in  $T_{syn} - T_R$ . The major parameter of IM is  $T_{syn}$  since it is determined as a gasification time averaged parameter. However energy assessment of syngas should also be expressed in reference to GC analysis.

Syngas characteristics determined by GC analysis are identical by using real gas and perfect gas treatments due to the treatment at a reference temperature of  $T_R = 20^\circ C$ . The similarity in RD with  $M_n$  is confirming the validity of IM. Energy of syngas is considerably different in IM and GC due to three basic reasons

- 1) Syngas temperature is rather higher than  $T_R = 20^\circ C$ . Therefore sensible energy is added with a higher  $E_{syn}$  in GC than the one in IM
- 2) Stored syngas sample for gas chromatography analysis is not the total decomposed fuel.

3) Fuel decomposition -gasification is not steady . Therefore syngas production is not at a constant rate and characteristics instead it is a transient procedure.

The major parameters are  $T_{syn}$ ,  $T_R$ ,  $E_{syn}$ ,  $W_I$ ,  $R_D$ ,  $M_n$  of syngas. Thermochemical decomposition of fuel is absolutely a solid function of fuel content. The conversion of C and H in fuel to  $CH_4$ ,  $CO$ ,  $CO_2$ , and  $H_2$  components of syngas describes the process. Furthermore the respective orders of  $CO/CO_2$  ratio and  $H_2/CH_4$  ratio of syngas indicate the path of thermochemical decomposition on which further analysis is vital for a complete understanding .

### References

- [1] Sanlisoy, A Carpinlioglu Ozdinc M ,”A review on plasma gasification for solid waste disposal” ,International Journal of Hydrogen Energy Vol 42 , 2017,pp 1361-1365.
- [2] Carpinlioglu Ozdinc M, Sanlisoy A, “Performance assessment of plasma gasification for waste to energy conversion: A methodology for thermodynamic analysis” International Journal of Hydrogen Energy Vol.43, 2018, pp 11493-11504.
- [3] Sanlisoy, A Carpinlioglu Ozdinc M “ Microwave Plasma Gasification of a Variety of Fuel for Syngas Production”Plasma Chemistry and Processing V:39,2019 Issue:5 pp: 1211-1225
- [4] Sanlisoy , A., An Experimental Investigation on Design and Performance of Plasma Gasification Systems, Ph. D thesis in Mechanical Engineering Department of Gaziantep University TURKEY ,2018
- [5] Carpinlioglu Ozdinc M , Design Construction and Performance Assessment of a Test Plant(Microwave Gasifier) “MCwgasifier” Using Plasma Gasification for Solid Waste-Energy Conversion in Laboratory Scale – An Experimental Case For Knowhow On Plasma Technology ,TUBITAK 115M389 Final Project Report ,2018
- [6] Carpinlioglu Ozdinc M “Perspectives on waste to energy conversion by microwave plasma gasification “Waste to resources 9 th International Symposium MBT ,MRF &Recycling Online Conference 18-20 May 2021 Hanover GERMANY
- [7] Bishoge, Obadia Kyetuza; Huang, Xinmei; Zhang, Lingling; et al” The adaptation of waste-to-energy technologies: towards the conver-sion of municipal solid waste into a renewable energy resource”Environmental Reviews Volume: 27 Issue: 4 ,2019 Pages: 435-446
- [8] Munir, M. T.; Mardon, I.; Al-Zuhair, S.; et al.” Plasma gasification of municipal solid waste for waste-to-value processing” Renewable & Sustainable Energy Reviews Volume: 116 ,2019 Article Number: 109461
- [9] Gimzauskaite, Dovile; Tamo-siunas, Andrius; Tuckute, Simona; et al” Treatment of diesel-contaminated soil using thermal water vapor arc plasma” Environmental Science and Pollution Rese-arch Volume: 27 ,2020 Issue: 1 Special Issue: SI Pages: 43-54
- [10] Mukherjee,C.; Denney, J.; Mbo-nimpa, E. G.;et al.” A review on municipal solid waste-to-energy trends in the USA” Renewable & Sustainable Energy Reviews Volume: 119 ,2020 Article Number: 109512
- [11] Gadzhiev, M. Kh.; Kulikov, Yu. M.; Son, E. E.; et al” Efficient Generator of Low-temperature Argon Plasma with an Expanding Channel of the Output” High Temperature Volume: 58 ,2020 Issue: 1 Pages: 12-20
- [12] Wimcompass Hobre Instrument 2018 General Information Wobbe Index and Calorimeters[https://www.hobre.com/files/products/Wobbe Index General Information.rev.1. pdf](https://www.hobre.com/files/products/Wobbe%20Index%20General%20Information.rev.1.pdf)
- [13] Neutrium, 2018 Wobbe Index [https:// neutrium.net/properties /wobbe-index/](https://neutrium.net/properties/wobbe-index/) 08.04.2018

# 3D Object Detection using Mobile Stereo R-CNN on Nvidia Jetson TX2

Mohamed K. Hussein  
Computer and Systems Engineering  
Department  
Ain Shams University  
Cairo, Egypt  
mohamed.khaled@eng.asu.edu.eg

Mahmoud I. Khalil  
Computer and Systems Engineering  
Department  
Ain Shams University  
Cairo, Egypt  
mahmoud.khalil@eng.asu.edu.eg

Bassem A. Abdullah  
Computer and Systems Engineering  
Department  
Ain Shams University  
Cairo, Egypt  
babdullah@eng.asu.edu.eg

**Abstract**—3D Object detection is one of the most important perception tasks needed by autonomous vehicles to detect different road agents like other vehicles, cyclists, and pedestrians which is essential for driving tasks like collision avoidance and path planning. In this paper, our work is focused on 3D Object Detection for car class from stereo images without LIDAR supervision neither during training nor during inference and the challenging task of running 3D Object Detection on an embedded target Nvidia Jetson TX2 by modifying Stereo R-CNN model and reducing the model size to approximately one third the size of the original model to be more suitable for embedded targets. Experiments on KITTI dataset showed that our model's inference time is 1.8 seconds and its' average precision for moderate car class is 17% on the test set. Our model decreases training and inference time by approximately 60% with a 13% drop on the test set which is an expected trade-off when decreasing the number of parameters inside the model.

**Keywords**—3D Object Detection, Stereo Vision, Autonomous Driving, Embedded Systems

## I. Introduction

3D Object Detection extends 2D object detection by also detecting pose and orientation as well as real-world dimensions of detected objects as shown in Figure I. This task is essential in autonomous driving for tasks like motion prediction and accurate object localization which is important for safe driving.

The most used sensors for 3D Object Detection are LIDAR, stereo camera, and mono camera. State of the art models heavily rely on LIDARs but LIDARs have the disadvantage of being very expensive which can cost around \$70000 compared to cameras which can cost less than \$1000. Also, LIDAR point clouds contain sparse information compared to dense information in images. The main advantage of LIDAR is accurate depth information which is essential for 3D Object Detection and cameras cannot provide the same level of accuracy and this is why there is a huge gap in average precision between LIDAR-based methods and camera-based methods. Camera-based methods are divided into stereo-based methods and methods relying on just one RGB image. Stereo-based methods perform better than mono-based methods because depth can be better estimated from stereo images.

While average precision is usually the metric used for comparing different methods, in this paper our work is focused more on the target platform and running 3D Object Detection on embedded target Nvidia Jetson TX2 with reduced memory usage and less inference time and our contribution is testing our proposed on KITTI [45] 3D Object Detection task using an embedded target while other methods use more powerful GPUs and sometimes use multiple GPUs for training and inference. For this purpose, we modified the

Stereo R-CNN [1] model which is a stereo-based end-to-end deep neural network that extends Faster R-CNN [2] and it does not rely on LIDAR supervision during training and inference. Nvidia Jetson TX2 has 256-core Nvidia Pascal GPU architecture with 256 Nvidia CUDA cores, Dual-Core Nvidia Denver 2 64-bit CPU and Quad-Core ARM Cortex-A57 MPCore, and 8 GB 128-bit LPDDR4 Memory.

In the next section, we briefly review different methods used for 3D Object Detection. In section III, we illustrate the modifications made to the Stereo R-CNN network to be more suitable for an embedded target. In section IV, we explain the experiment setup and what was done and report the results achieved and compare it to the original model based on the evaluation on KITTI validation set and test set. In section V, we conclude this paper with a summary of our work and contribution.

## II. Related Work

### A. LIDAR Based Methods

LIDAR-based methods can be mostly classified into two categories depending on how the point cloud is represented. **Grid-based methods** depend on point cloud voxelization like [3]-[8] or projecting point clouds to 2D grids [9] and then these grids are processed by 2D or 3D CNN. **Point-based methods** process raw point clouds directly like [10]-[15]. Also, some methods use both representations like [16]-[18] and [46]. There are other categories like [19] which uses a graph representation for point clouds. At the time of writing this paper (April 2021), the top LIDAR method on KITTI 3D benchmark is SE-SSD [46] which has an average precision of 82.54% in the car moderate class. This is also the top method regardless of approach.

### B. LIDAR + Image Fusion Based Methods

Fusion methods can be mostly classified into three categories. **Early Fusion** in which inputs from camera and LIDAR are fused very early in the input of the deep neural network like [20]. **Late Fusion** in which there is a separate pipeline for Camera and LIDAR and they are fused at a later stage in the model like [21]-[27]. **Deep Fusion** in which inputs are fused multiple times deep inside the model like the leading work in 3D Object Detection MV3D like [28]-[30]. At the time of writing this paper, the top fusion method on KITTI 3D benchmark is CLOCs [27] which has an average precision of 80.67% in the car moderate class.

### C. Stereo Images Based Methods

Stereo-based methods can be mostly classified into two categories. **LIDAR supervised methods** that need LIDAR input for training and use only stereo images during inference like [31]-[38]. Pseudo-LIDAR was proposed in [31] to reduce the gap between 3D object detection using LIDAR and stereo camera. It relies on building point clouds from depth maps and



Figure I. 2D Object Detection (top) and 3D Object Detection (bottom)

feeding the generated point cloud to a LIDAR-based 3D Object Detector and this approach achieved the best average precision on KITTI benchmark when it was published. Pseudo-LIDAR can be used with mono or stereo images by just changing the depth estimation module which creates the depth map that is later transformed to a pseudo point cloud. In [36], Pseudo-LIDAR++ was proposed which builds over [31] by enhancing depth estimation which is the first step to produce pseudo point cloud. In [34], end-to-end Pseudo-LIDAR was proposed since previous methods trained two different blocks for depth estimation and 3D object detection separately. Deep Stereo Geometry Network for 3D object detection was proposed in [32] which is an end-to-end network relying on constructing a 3D geometric volume and plane-sweep volume for 2D features then use the 3D geometric volume for 3D Object detection and the plane-sweep volume for Depth Estimation simultaneously. In [33], confidence-guided 3D object detection is proposed by doing depth estimation separately for foreground and background pixels and giving a confidence estimation for each pixel which is used with the generated point cloud as input for a 3D Object Detector. In [35], Zoomnet was proposed and in this approach, the 2D Region of Interests were resized to have the same resolution so that near and far objects are analyzed with the same resolution and instance point cloud is generated from the depth estimation for each instance bounding box. In [37], Object Centric stereo matching is proposed trying to enhance the stereo matching problem for 3D object detection because other approaches used depth estimation networks with depth maps as the main output not point clouds so in the approach the instance segmentation is done, and object-centric instance point clouds are generated to enhance the produced point clouds which are then fed to a 3D object detector. In [38], a Continuous Disparity Network (CDN) is proposed with a Wasserstein distance-based loss function to enhance disparity estimation and this CDN can be used with any stereo 3D Object Detector that relies on disparity estimation like DSGN [32]. **No LIDAR supervision methods** only use stereo images in both training and inference like [1] and [39]-[41]. Stereo R-CNN [1] is the baseline model for our work so it will be discussed later in section III. In [39], the images are used to generate a semantic map with 2D bounding boxes and a

disparity image which is then projected to a grid which is then used to estimate the 3D bounding box. In [40], Instance Depth Aware (IDA) 3D Object Detection is proposed in which stereo region proposal network is used to get 2D bounding boxes then IDA module estimates the center of the 3D bounding box and the region proposals are used to determine position and orientation of the 3D bounding box. In [41], Disp R-CNN is proposed in which stereo pair is used as input to Stereo Mask R-CNN network to produce instance masks for objects of interest then instance disparity is generated which is used to generate instance point clouds which are fed to a 3D Object Detector.

Stereo methods that rely on the concept of generating point clouds perform better than methods that use images directly for 3D Object Detection but most of them require the presence of ground truth point clouds during training and needs a lot of computation power as there are two steps in the process, first you need to generate a pseudo point cloud then feed it to a point cloud-based 3D Object Detector.

At the time of writing this paper, the top stereo method on KITTI 3D benchmark is CDN-DSGN [38] which has an average precision of 54.22% in the car moderate class.

#### D. Mono Images Based Methods

Monocular 3D Object Detection is the most challenging method due to the complete lack of depth information and therefore there is a huge performance gap between monocular methods and other methods. At the time of writing this paper, the top monocular method [42] on KITTI 3D benchmark has an average precision of 12.72% in car moderate class which is a video-based method that uses temporal cues and kinematics to improve localization accuracy.

### III. Mobile Stereo R-CNN

Mobile Stereo R-CNN as shown in Figure II is based on Stereo R-CNN which is an end-to-end deep neural network for 3D Object Detection from stereo images without the supervision of LIDAR in neither training nor inference. Our main target was running 3D Object detection on an embedded target like NVIDIA Jetson TX2. So first we analyzed the building blocks of Stereo R-CNN to identify bottlenecks that

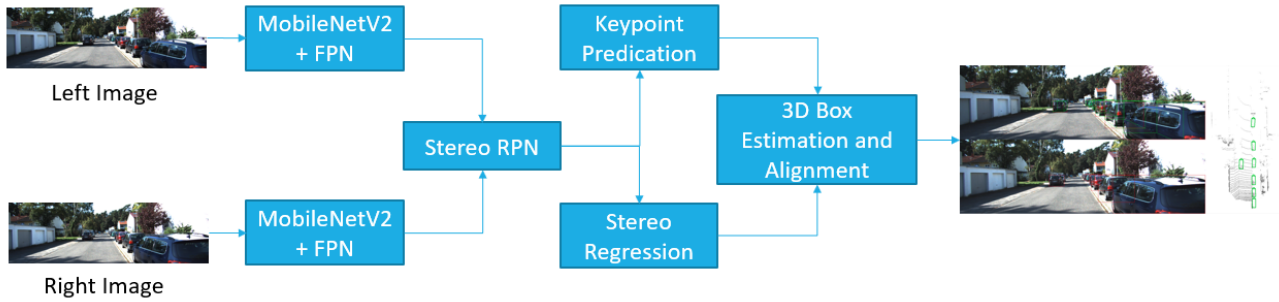


Figure II. Mobile Stereo RCNN Network Architecture

can be modified to enhance the training time, inference time, and memory footprint and we decided to change the backbone network from ResNet101-FPN to MobileNetV2-FPN. The new model size is approximately third the size of the original model and inference time is reduced by approximately 60% with a minor drop of 6% in average precision on moderate car class on KITTI validation set and a drop of 13% in average precision on moderate car class on KITTI test set. First, we review the building blocks of Stereo R-CNN then discuss the backbone replacement.

### A. Stereo R-CNN

Stereo R-CNN is based on Faster R-CNN. It takes as input two stereo images resized so that the shortest side length is 600px.

**Backbone Network.** ResNet101-FPN is used as the backbone network for Stereo R-CNN which takes as input the input stereo images and produces 2D feature maps. Both backbone networks for right and left images use shared weights. The feature maps are then concatenated and used as input to Stereo Region Proposal Network (Stereo RPN).

**Stereo RPN.** Stereo RPN has two branches, a branch for classification that is responsible to measure the objectness of each anchor and a branch for regressing the box offsets. Stereo RPN produces left and right Regions of Interests (RoIs).

RoI alignment is done on both left and right RoIs separately. Left feature maps with RoIs are used as input for keypoint prediction and concatenated left and right features with RoIs are used as input for stereo regression.

**Keypoint Prediction.** In this part of the model 3D Semantic keypoints are predicted to enhance the 3D box estimation using convolution layers.

**Stereo Regression.** In this part of the model the following tasks are done: object classification, regression of 2D bounding boxes, regression of dimension, and regression of viewpoint angle.

**3D Box Estimation.** Outputs from stereo regression and keypoint prediction blocks are then taken as input for 3D box estimation and alignment to produce final 3D boxes of 3D object detection.

### B. Backbone Replacement

The new backbone is based on MobileNetV2 [43] which is more suitable for mobile and embedded targets than ResNet101. We also use a feature pyramid network with MobileNetV2 to maintain the region proposal network as it is in the original Stereo R-CNN implementation. We divide the MobileNetV2 into 5 layers as shown in table I. Each entry in table I is in the format (scale ratio with respect to input image

\* the number of channels) these feature maps are then used with a feature pyramid network to produce the output of the base model with 128 channels with the following dimension ratios 1/4, 1/8, 1/16, 1/32, and 1/64 relative to the input dimension. Five anchor scales 32, 64, 128, 126, 512, and three ratios 0.5, 1, 2 are used in the network.

Layer 0 consists of a 2D convolution operation and 2 bottleneck operations. Layer 1 consists of 1 bottleneck operation. Layer 2 consists of 3 bottleneck operations. Layer 3 consists of 7 bottleneck operations. Layer 4 consists of 4 bottleneck operations followed by a 2D convolution to produce the final feature map. The reason for choosing these layers was to match the lateral connections done to the feature pyramid network of the original model.

TABLE I. MobileNetV2-FPN Backbone

Layer	Input	Output
0	1 x 3	1/4 x 24
1	1/4 x 24	1/4 x 24
2	1/4 x 24	1/8 x 32
3	1/8 x 32	1/16 x 96
4	1/16 x 96	1/32 x 1280

## IV. Experiment and Results

We split the training set of KITTI to 3712 frames for training and 3769 frames for validation using the data split done by [44] which is also the same split used by Stereo R-CNN so that we can compare our results on the same validation set with the base model. Data augmentation is done for the 3712 frames of the training set using the same approach as Stereo R-CNN with stereo flipping so the final training set size is 6478 frames.

### A. Training

Training is performed on Nvidia GeForce GTX 1080Ti GPU for 11 epochs for 14 hours compared to 33 hours using the original Stereo R-CNN model. Batch size equals 1 and the network is trained using SGD with a momentum of 0.9 and weight decay of 0.0005. The learning rate equals 0.001 and is reduced by 0.1 every 5 epochs. The model is trained end-to-end but the backbone layers 0 to 2 were frozen during training to maintain the pre-trained weights of MobileNetV2.

### B. Inference

The inference is done on both Nvidia GeForce GTX 1080Ti GPU and Nvidia Jetson TX2. Inference time on Nvidia GeForce GTX 1080Ti GPU is 0.16 seconds compared to 0.36 seconds using the original Stereo R-CNN model. Inference time on Nvidia Jetson TX2 is 1.8 seconds compared to 4.7 seconds using the original Stereo R-CNN model.



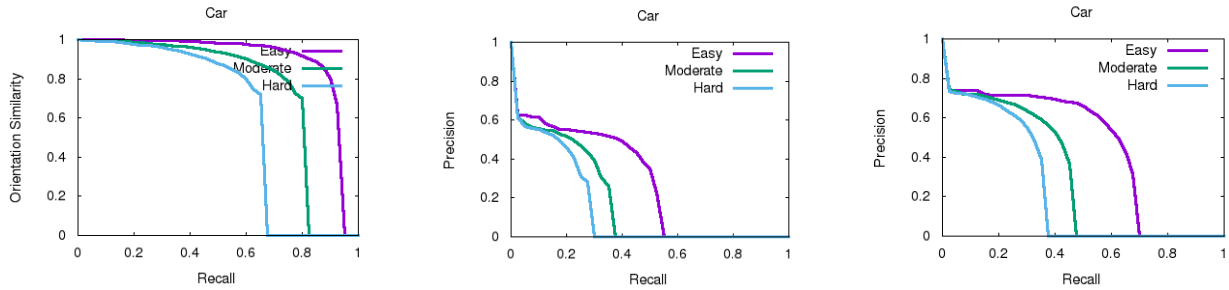


Figure III. Precision-Recall curves for Orientation estimation, 3D object detection, and Bird's eye view detection respectively

TABLE II. Mobile Stereo-RCNN vs Stereo R-CNN on KITTI Validation Set

Model	AP (Easy)	AP (Medium)	AP (Hard)	Inference Time on Jetson TX2	Model Size (Number of Parameters)
Stereo R-CNN	54.11%	36.69%	31.07%	4.7 seconds	Approx. 105M
Mobile Stereo R-CNN	44.83%	30.77%	25.25%	1.8 seconds	Approx. 35M

TABLE III. Mobile Stereo-RCNN vs Stereo R-CNN on KITTI Test Set

Model	Stereo R-CNN			Mobile Stereo R-CNN		
	Benchmark	Easy	Moderate	Hard	Easy	Moderate
Car (Detection)	93.98%	85.98%	71.25%	90.08%	76.73%	62.23%
Car (3D Detection)	47.58%	30.23%	23.72%	26.97%	17.04%	13.26%
Car (Bird's Eye View)	61.92%	41.31%	33.42%	44.51%	28.78%	22.30%

### C. Results

The new model as shown by the experiments done reduces model size and inference time with a minor drop in average precision. Table II shows the comparison between the performance of Stereo R-CNN and Mobile Stereo R-CNN on Nvidia Jetson TX2 on KITTI validation set. Table III shows a comparison between the results of Mobile Stereo R-CNN and Stereo R-CNN from KITTI test set. Figure III shows the precision-recall curves from KITTI benchmark for Orientation estimation, 3D object detection, and Bird's eye view detection respectively. Our model average precision is less than the original model with a 13% drop in moderate car class which is an expected trade-off after decreasing the model size and its' number of parameters by replacing ResNet101 with MobileNetV2 which is more suitable for running on embedded platforms.

### Acknowledgment

We would like to thank eJad (<http://www.ejad.com.eg>) for their technical and logistics support for this work.

### V. Conclusion

In this paper, our main contribution is running 3D Object Detection on embedded target Nvidia Jetson TX 2 using only stereo images without LIDAR supervision. we base our work on Stereo R-CNN and replace its' ResNet101-FPN backbone with a new MobileNetV2-FPN backbone that is more suitable for embedded devices and with this we achieved a model size approximately third of the original model and inference time that is approximately 40% of the original model with a minor drop of 13% in average precision.

### References

- [1] P. Li, X. Chen, and S. Shen, "Stereo R-CNN based 3d object detection for autonomous driving," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 2019.
- [2] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," in Advances in Neural Information Processing Systems, vol. 28, 2015.
- [3] Y. Yan, Y. Mao, and B. Li, "SECOND: Sparsely embedded convolutional detection," Sensors, vol. 18, p. 3337, Oct. 2018.
- [4] J. Lehner, A. Mitterecker, T. Adler, M. Hofmarcher, B. Nessler, and S. Hochreiter, "Patch refinement – localized 3d object detection," 2019.
- [5] Z. Liu, X. Zhao, T. Huang, R. Hu, Y. Zhou, and X. Bai, "TANET: Robust 3d object detection from point clouds with triple attention," 2019.
- [6] B. Wang, J. An, and J. Cao, "Voxel-FPN: multi-scale voxel feature aggregation in 3d object detection from point clouds," 2019.
- [7] Y. Ye, H. Chen, C. Zhang, X. Hao, and Z. Zhang, "SARNET: Shape attention regional proposal network for LIDAR-based 3d object detection," Neurocomputing, vol. 379, pp. 53–63, Feb. 2020.
- [8] H. Yi, S. Shi, M. Ding, J. Sun, K. Xu, H. Zhou, Z. Wang, S. Li, and G. Wang, "Segvoxelnet: Exploring semantic context and depth-aware features for 3d vehicle detection from point cloud," 2020.
- [9] Y. Zeng, Y. Hu, S. Liu, J. Ye, Y. Han, X. Li, and N. Sun, "RT3d: Real-time 3-d vehicle detection in LiDAR point cloud for autonomous driving," IEEE Robotics and Automation Letters, vol. 3, pp. 3434–3440, Oct. 2018.
- [10] C. R. Qi, L. Yi, H. Su, and L. J. Guibas, "Pointnet++: Deep hierarchical feature learning on point sets in a metric space," 2017.
- [11] C. R. Qi, H. Su, K. Mo, and L. J. Guibas, "Pointnet: Deep learning on point sets for 3d classification and segmentation," in Proceedings of the Conference on Computer Vision and Pattern Recognition (CVPR), July 2017.
- [12] A. H. Lang, S. Vora, H. Caesar, L. Zhou, J. Yang, and O. Beijbom, "Pointpillars: Fast encoders for object detection from point clouds," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 2019.
- [13] S. Shi, X. Wang, and H. Li, "PointRCNN: 3d object proposal generation and detection from point cloud," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 2019.
- [14] Z. Yang, Y. Sun, S. Liu, X. Shen, and J. Jia, "Std: Sparse-to-dense 3d object detector for point cloud," in Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), October 2019.
- [15] Z. Yang, Y. Sun, S. Liu, and J. Jia, "3dssd: Point-based 3d single stage object detector," 2020.

- [16] Y. Chen, S. Liu, X. Shen, and J. Jia, "Fast point R-CNN," in Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), October 2019.
- [17] C. He, H. Zeng, J. Huang, X.-S. Hua, and L. Zhang, "Structure aware single-stage 3d object detection from point cloud," in IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 2020.
- [18] S. Shi, C. Guo, L. Jiang, Z. Wang, J. Shi, X. Wang, and H. Li, "PV-RCNN: Point-voxel feature set abstraction for 3d object detection," in IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 2020.
- [19] W. Shi and R. Rajkumar, "Point-GNN: Graph neural network for 3d object detection in a point cloud," in IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 2020.
- [20] M. Simon, K. Amende, A. Kraus, J. Honer, T. Samann, H. Kaulbersch, S. Milz, and H. Michael Gross, "Complexer-yolo: Real-time 3d object detection and tracking on semantic point clouds," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops, June 2019.
- [21] X. Du, M. H. A. J. au2, S. Karaman, and D. Rus, "A general pipeline for 3d detection of vehicles," 2018.
- [22] C. R. Qi, W. Liu, C. Wu, H. Su, and L. J. Guibas, "Frustum pointnets for 3d object detection from RGB-d data," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 2018.
- [23] M. Liang, B. Yang, Y. Chen, R. Hu, and R. Urtasun, "Multi-task multi-sensor fusion for 3d object detection," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 2019.
- [24] Z. Wang and K. Jia, "Frustum convnet: Sliding frustums to aggregate local point-wise features for amodal 3d object detection," 2019.
- [25] S. Vora, A. H. Lang, B. Helou, and O. Beijbom, "Pointpainting: Sequential fusion for 3d object detection," in IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 2020.
- [26] J. H. Yoo, Y. Kim, J. Kim, and J. W. Choi, "3d-cvf: Generating joint camera and lidar features using cross-view spatial feature fusion for 3d object detection," Lecture Notes in Computer Science, p. 720–736, 2020.
- [27] S. Pang, D. Morris, and H. Radha, "Clocs: Camera-lidar object candidates fusion for 3d object detection," 2020.
- [28] X. Chen, H. Ma, J. Wan, B. Li, and T. Xia, "Multi-view 3d object detection network for autonomous driving," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 2017.
- [29] J. Ku, M. Mozifian, J. Lee, A. Harakeh, and S. Waslander, "Joint 3d proposal generation and object detection from view aggregation," 2018.
- [30] M. Liang, B. Yang, S. Wang, and R. Urtasun, "Deep continuous fusion for multi-sensor 3d object detection," in Proceedings of the European Conference on Computer Vision (ECCV), September 2018.
- [31] Y. Wang, W.-L. Chao, D. Garg, B. Hariharan, M. Campbell, and K. Q. Weinberger, "Pseudo-lidar from visual depth estimation: Bridging the gap in 3d object detection for autonomous driving," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 2019.
- [32] Y. Chen, S. Liu, X. Shen, and J. Jia, "DSGN: Deep stereo geometry network for 3d object detection," in IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 2020.
- [33] C. Li, J. Ku, and S. L. Waslander, "Confidence guided stereo 3d object detection with split depth estimation," 2020.
- [34] R. Qian, D. Garg, Y. Wang, Y. You, S. Belongie, B. Hariharan, M. Campbell, K. Q. Weinberger, and W.-L. Chao, "End-to-end pseudo-lidar for image-based 3d object detection," in IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 2020.
- [35] Z. Xu, W. Zhang, X. Ye, X. Tan, W. Yang, S. Wen, E. Ding, A. Meng, and L. Huang, "Zoomnet: Part-aware adaptive zooming neural network for 3d object detection," 2020.
- [36] Y. You, Y. Wang, W.-L. Chao, D. Garg, G. Pleiss, B. Hariharan, M. Campbell, and K. Q. Weinberger, "Pseudo-lidar++: Accurate depth for 3d object detection in autonomous driving," 2020.
- [37] A. D. Pon, J. Ku, C. Li, and S. L. Waslander, "Object-centric stereo matching for 3d object detection," in 2020 IEEE International Conference on Robotics and Automation (ICRA), IEEE, May 2020.
- [38] D. Garg, Y. Wang, B. Hariharan, M. Campbell, K. Weinberger, and W.-L. Chao, "Wasserstein distances for stereo disparity estimation," in NeurIPS, 2020.
- [39] H. Konigshof, N. O. Salscheider, and C. Stiller, "Realtime 3d object detection for automated driving using stereo vision and semantic information," in 2019 IEEE Intelligent Transportation Systems Conference (ITSC), IEEE, Oct. 2019.
- [40] W. Peng, H. Pan, H. Liu, and Y. Sun, "Ida-3d: Instance-depth-aware 3d object detection from stereo vision for autonomous driving," in IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 2020.
- [41] J. Sun, L. Chen, Y. Xie, S. Zhang, Q. Jiang, X. Zhou, and H. Bao, "Disp R-CNN: Stereo 3d object detection via shape prior guided instance disparity estimation," in IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 2020.
- [42] G. Brazil, G. Pons-Moll, X. Liu, and B. Schiele, "Kinematic 3d object detection in monocular video," in Computer Vision – ECCV 2020, pp. 135–152, Springer International Publishing, 2020.
- [43] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "Mobilenetv2: Inverted residuals and linear bottlenecks," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 2018.
- [44] X. Chen, K. Kundu, Y. Zhu, H. Ma, S. Fidler, and R. Urtasun, "3d object proposals using stereo imagery for accurate object class detection," 2017.
- [45] Geiger, P. Lenz, and R. Urtasun, "Are we ready for autonomous driving? the kitti vision benchmark suite," in Conference on Computer Vision and Pattern Recognition (CVPR), 2012.
- [46] Z. Li, Y. Yao, Z. Quan, W. Yang, and J. Xie, "Sienet: Spatial information enhancement network for 3d object detection from point cloud," 2021.

# Investigation of Algorithmic Architecture Design Method by Using Digital Technology to Increase Flexibility in Design Process

Aida Ghiaseddin  
Department of Architecture  
Islamic Azad University  
Tehran, Iran  
[st\\_a\\_ghiaseddin@azad.ac.ir](mailto:st_a_ghiaseddin@azad.ac.ir)

Javad Eiraji  
Faculty of Architecture and Design,  
Eskisehir Technical University  
Eskisehir, Turkey  
[javadeiraji@eskisehir.edu.tr](mailto:javadeiraji@eskisehir.edu.tr)

**Abstract**—In metropolitan cities architectures projects are full of diversity and innovation. By observing these global architectural projects causes new ideas and by this way we can understand these architectural diversity. In addition, when these architectural varieties increases, for using conventional tools and methods in designing architectural projects we need to provide new methods and tools. One of these tools and methods is designing by an algorithmic method. This algorithmic architecture method quickly allows the designer to create various designs and causes precise control over the project during the architectural process. This paper is a qualitative research which seeks to evaluate algorithmic design as a powerful tool in digital architecture efficiency and implementing flexible and creative ideas. The aim of this study, is to explore an approach that increases the designers' creativities and provides a framework for the basic requirements of the architectural design.

**Keywords**—Algorithmic Design Process, Performance-Based Design, Parametric Architecture, Computer-Aided Design, Flexible Design

## I. Introduction

Today architectural design are more flexible by using new software, which help us to produce significant number of choices during the design process.

Since the mid-1990s, architectural approaches started to focus on parametric graphic design in architecture and it was increased considerably. During this period, software such as Grasshopper and Generative Component(GC) was developed with advanced features and then evolved rapidly. In 2000, the applications of the parametric techniques in designing buildings had matured so that an increasing number of spectacular monuments proposed and built by using the parametric design method. In terms of research and education, the AA in London, LAAC in Spain, MIT, and the University of Colombia in the US and, Hyper body lab in the Delft University of Technology turned into an origin for the next generation of parameter designers. Meanwhile, the international conferences on computational design, including ACADIA (Association for Computer Aided Design in Architecture) , ASCAAD (Arab Society for Computer Aided Architectural Design), CAADRIA (Computer-Aided Architectural Design Research in Asia) and eCAADe (Education and research in Computer Aided Architectural Design in Europe), publish many proceedings , which is a critical subject founded in these research halls since that time. Recently, many other workshops and conferences were held to progress the parametric design objectives [1].

The term "algotecture" was taken from the algorithmic architecture and represents algorithms for architectural design. It has become widespread in architectural design over the past few decades. The parametric instruments work based on algorithms which can apply exact control over geometry of the design throughout the design process. The flexibility and responsiveness of these instruments to the design changes have result in the usefulness and applicability of the parametric models, particularly in designing complex and unique models [2].

## II. Literature Review

According to Patrick Schumacher, a successful design employs technologies and instruments, which help the designer proceed with his design. Regarding to algorithmic and parametric architecture, he states that the architecture should move from a single-layer system and an application for the design editions toward a multi-layer and yet, consistent and continuous design of the multisystem such as envelope, structure, and internal subsection. The application of any design operation on a multisystem must be correlated with other components of the system and influence them.

In his well-known paper, Schumacher mentions parametric design as a deep style and introduces the word parametricism as a newly emerging approach for articulating architecture, which provides an acceptable validity [3].

Although some degree of compatibility of style with parametric architecture yields an appropriate method for generating free-form and complex-geometry designs, parametric design is not limited to any formal language. It is not a style but, instead, it is a design method rather than a style.

Meredith points to the threat against parametric design, which can quickly turn into a visual aestheticism integrated with a tendency toward formalism [4].

Similarly, Mousavi states the parametric design as a style that deprives itself of the limitations of the external parameters and improves the independence of the architectural forms. However, it cannot go beyond the new ways of shaping the materials for producing unexpected spaces [5] [6].

## III. Theoretical fundamentals

Parametric, as a term or method, is rooted in the mathematical concept of parametric equations. Its design program was observed first in the works of the Italian architect Luigi Moretti in the 1950s [7]. The fundamental idea of parametric design is that the values of different variables usually vary associatively based on different input parameters.

The conception of this design is originated from production design that its primary purpose is to help human designers to discover the design space by computational instruments [8]. Thus this purpose can be realized by computers through fast sampling. Parametric design systems discuss as production instruments in architectural design and the parametric instruments are deployed algorithmically so apply more computational control over the design geometry.

The role of parametric modeling addresses as a production design instrument in architecture. Parametric design is a computational method that can act as a productive (generative) method and an analysis. Besides, it has recently received good acceptability in terms of practicality, research, and education. There are some debates about the limitations of the parametric systems as an exploratory instrument, mainly addressing its role in architectural design, flexibility, and design complexity. According to its algorithmic basis and its potential for expanding the design discovery space by changing the algorithm variables, i.e., parameters, the parametric design can be classified as the third class of the production systems [9].

Algorithmic thinking and algorithmic design correlate with the concept of product design. Terzidis argues that the inductive strategy of the algorithms can discover the production processes or simulate complex phenomena. Algorithms can be assumed as the suffix of the human brain and might facilitate mutation in some regions of the unpredictable potential [10].

### A. Parametric Design Systems

The term parametricism introduced by (P. Schumacher) for the first time and later, has been described more comprehensively as a combination of design concepts, which provides a new and complex discipline based on the fundamental principles.

Parametric systems work due to algorithmic rules. An algorithm is a limited set of instructions for achieving a specific aim. An algorithm takes a value or a group of values as the input, displays several countable stages, which convert or change the input, and finally generates one or multiple values as the output [11].

A parameter is the value or measurement of a variable that can vary. Every object (thing) in a parametric system might have some specific rules. When a parameter changes, other parameters will be adapted to it automatically [12].

Parametric design usually uses complex building molds, energy and structural optimization, other repetitive works, and design sampling. As a novel digital design method, parametric design is entirely different from CAD/CAM because these algorithmic features are rule-based (Fig I).

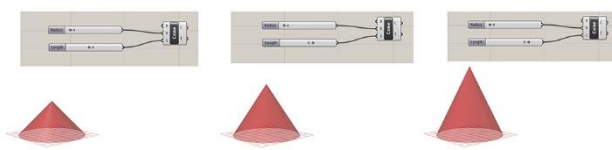


Figure I. Simplified examples of parametric variations generated in the same system .[1]

In terms of computations, there is no difference between algorithmic and parametric systems. Algorithms pre-supposedly work on parameters, and the main component of a parametric system is the algorithm itself, which is called a design or definition. Although, unlike algorithmic design, parametric systems emphasize explicit and direct manipulation of the parameter values to change and modify the design artifact.

The significant advantage of parametric modeling is that it allows us to change the parameters at any level of the design process [13].

So far, various parametric modeling techniques developed for visual purposes (e.g., finding form) and other functional or performance-related purposes. Figure II shows some of the usual parametric modeling techniques used for form-finding, e.g., repetition (Fig II – left) and division (Fig II – right). Other similar methods may include tiling and like that (Fig II).

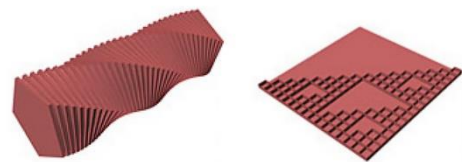


Figure II. Examples of parametric modeling techniques for form finding: Repetition (Left) and Subdivision (Right). [1]

Most software hold up free-form modeling, yet they have a scripting plug-in (add-on), enabling the designers to create the rule algorithms directly and more freely. Some of these software packages are introduced in short in this division. Rhino and Grasshopper are some of the most commonly known parametric instruments, particularly in architecture. However, Digital Project (DP) and Generative Component (GC) are more suitable for large projects with multiplex and geometric associations [13].

DP is a compelling software package that can effectively shoulder geometric and complicated parameters, making it an ideal choice for sizeable parametric design projects.

Rhino is an independent and NURBS-based instrument invented by Robert McNeil. Since 1990, it uses widely in different subject, including architecture, industrial design, jewelry design, automotive, and marine design. Furthermore, Grasshopper is a rule algorithm editor with a graphic interface incorporated into Rhino as a scripting plug-in. This structure makes specific definition files that link to the main parametric model in Rhino.

Rhino is commonly used as a production instrument, instead of a correction (modification) instrument, in the parametric design process. Compared to other parametric software, Rhino and Grasshopper are now widely used in practicing and education. Such a wide use attributes to the simplicity of its function as a visual programming instrument.

A significant difference between parametric design and other standard computer design methods is the ruleset converted into the main design elements and procedures [14].

In this technique, through the parametric design procedure, the architects can return to any stage of the design

process to change the parameters or revise the rules to modify the design procedure for different purposes or perform different tests and experiments. Such flexibility allows the architects to keep the design opened (Fig III).

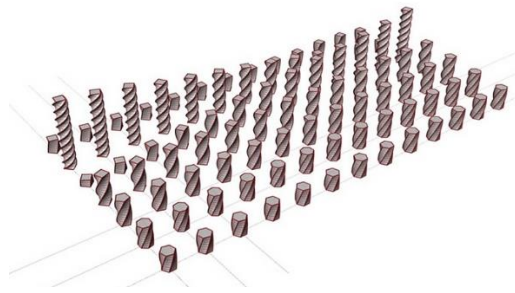


Figure III. A large number of form variations generated in a parametric design system. [1]

In parametric design, once the rule algorithms get created, many design choices will be generated (Fig III). This design sampling can expand the design abilities considerably and extend the designer's thinking. Additionally, the designers do not need to determine any solution early [15].

This feature allows keeping the maximum potential in the process. Parametric design is not only a new design instrument, it's a new method of design thinking [16].

**B. Describing Design based on Parametric Logic**

Parametric modeling as a design integration method diverges the design space to discover many kinds of similar parametric models. Parametric modeling can help a broader region for design discovery. A change in a parameter would cause a concurrent change in the form. While memorizing the basic coherence of the design, it applies the changes to the form.

In Figure IV the multiple geometric arrangements of the British Petrol Headquarters in Sunbury presented by Adams Kara Taylor, Shows the creative discovery of the roof structure is based on a parametric approach, which takes into consideration both aesthetically and structurally.

The Parametric model turn into a controlled environment in design discovery which searches for a greater design (Fig IV).

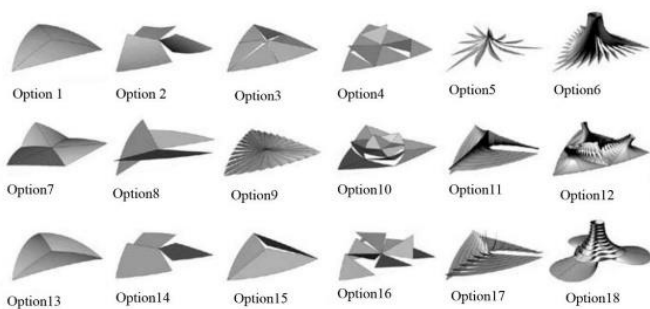


Figure IV. British Petrol Headquarters in Sunbury. [6]

Performative design principles can be integrated either early in the design process where design concepts and main geometry are being worked out, or later on during detail design where performance boost of systems is carried out. The Aviva Stadium in Dublin, Ireland (Figure V) and the façade of the kilden performing Arts Center in Kristiansand,

Norway are two works that are exemplary of these two approaches (Fig V).

During the design process, the architects ultimately were driving the overall form and cladding of the building and the engineers driving the structural member sizing / positioning. On the architectural side, certain form explorations were being made in response to certain criteria such as concourse width requirements, floor area ratios, or simply beautifying the shape on the engineering side were the structure of the roof trusses and cladding system designed as a rain screen consisting of inter-locking louvers. A single parametric model was shared between the architectural and engineering offices, which acted both as a design tool and a coordination platform. This allowed the integration of the design processes of the form, structure and façade, allowing fast response to design changes. Analysis tools were coupled with the parametric model and provided quick analytical reaction to the geometry. The sharing of the parametric model across the other design members and the fully integration of the engineering applications could be realized the benefits of a parametric approach (Fig VI).

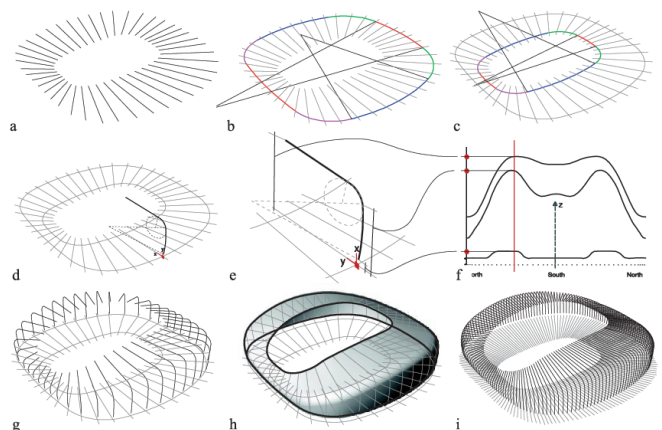
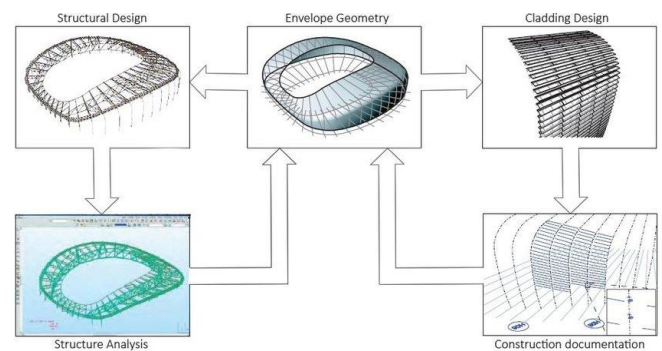


Figure V. Process of parametric design of Aviva Stadium in Dublin, Ireland. [6]

Figure VII shows the Kilden Performing Arts Center in Kristiansand, Norway, as another case of parametric modeling designed by ALA Architects and engineered by Design-to-Production Company. In this design, cantilevering up towards the waterfront, the timber façade intersects with a vertical glass and steel facade in both interior and exterior parts.



Figure VI. Aviva Stadium in Dublin, Ireland.[6]

The geometry of the façade is a ruled surface that spans between a straight upper and a curved lower edge. Here, a parametric system was used not during the form-finding process of the curvilinear roof, but during detail design for the parametric optimization of form and performance (Fig VII).



Figure VII. Model photo of Kilden Performing Arts Center. [6]

### C. Design Reports in Parametric Modelling

Several perspectives of the design space was applied to parametric design instruments. On one hand, a 3D model perspective demonstrates the geometric illustration (analogue); on the other hand, an editor allows the designer to encode the algorithm, i.e., schema. This editor is in either text or graphic forms. The visual schema editors (e.g., Grasshopper 3D plug-in by McNeil and Generative Component by Bentley) are now widely used due to the low level of technical knowledge required in programming languages.

### D. Details of Geometric Parameters Design

Another misunderstanding is that parametric design is the only way to generate complex geometry. even before computations, Complex geometry in architecture, has been an applied field in architectural design, as can be seen in the works of Frei Otto, Jorn Utzon, Pier Luigi Nervi, Felix Candela and Antonio Gaudi.

Each of these architects found creative techniques for overcoming the complexities of the free-form geometry in their works. For example, hanging the chain models was the physical discovery instrument for Antonio Gaudi during the form-finding process. However, this procedure was not fully efficient and smooth [17].

Another similar example is the Sydney Opera House by Jorn Utzon. It is a competitive project that could win the first-rank prize in 1957. Today, this monument is known as a masterpiece that catches the attention of many engineers and architects. However, the design discovery procedure has resulted in many differences at the time of its construction [18].

The geometry of the roof structure was not primarily defined and was considered unusable at first [19].

Over five years from designing the concept, the engineers and architects had to change the roof into an appropriate, appropriate form that allows the use of a unit mold and consequently, a unit bending during the construction stages [20].

This pattern by Gaudi and Utzon indicates a type of geometry that can reply to computational approaches and especially, parametric systems of performance.

## IV. Conclusion

The creation and evolution of algorithmic instruments enable the designer to create a new and debatable structure during the architectural design process and in addition, to answer to the environmental and physical needs using the new methods and proportionate to the design context. This approach is required when the structure and fundamental needs of the design have some complexities and covered layers, and the designer, despite mastership in designing, cannot manage and organize them. Thus, by providing different instruments and methods, parametric architecture can help the designers to control the design process by optimizing the project, bringing the time, financial, and human resources under the designer's control. As can be seen in the dead cubes scenario, the designer has performed using the algorithmic environment and instruments to respond to the design's fundamental requirements and have a novel method for designing residential buildings. This has guided the designer to employ the algorithmic instruments to create diverse and exciting spatial sequences aligned with his believes during the architectural design process for residential buildings. The notable point is the existence of different alternatives at the final presentation of the project, which indicates the considerable capabilities of algorithmic architecture in the design process and accomplishment of the process.

What is learned from investigating the parametric architecture design method by using digital technology aimed to improve flexibility is that a mutual and dynamic relationship between the components can cause maturity and evolution of the design. The digital design process improves the designer's creativity and mind power and causes his/her mind to work more dynamically and precisely; besides, it increases the designer's mental archive to the numerousness and diversity of the ideas and architectural structures. In short, the algorithmic architecture employs the designer's creativity as a computer that does not merely take the quantities into account and engages it in the design process. This approach aligns architecture with some imaginations that were far from being brought into action until a short while ago. However, it is obvious that this method will be effective when, in terms of qualitative criteria, the designer's mind continuously evaluates and controls the design process as an intelligent supervisor.

### References

- [1] Ning Gu, Rongrong Yu, and Peiman Amini Behbahani, "Parametric Design: Theoretical Development and Algorithmic Foundation for Design Generation in Architecture", Handbook of the Mathematics of the Arts and Sciences, Springer, 2018.
- [2] Moghtadanezhad, Pashaei, mehdi, Sevda, "Investigation of the Impact of Parametric Architectural Design Process Based on Algorithmic Design, A New Method in Digital Architectural Design for Achieving Sustainable Architectural Goals" 3rd international conference on Modern research in civil engineering architectural and urban development, Berlin-Germany, 2016.
- [3] Patrick Schumacher, "Parametricism: Rethinking Architectures Agenda for the 21st century (architectural design)" Academy Press, 2016.
- [4] Michael Meredith, AGU, Mutsuro Sasaki, P .ART, Design to production, Aranda, "From control to design: Parametric/algorithmic architecture", Actar; English edition, 2008.
- [5] Farshid Moussavi, "Parametric software is no substitute for parametric thinking", The architectural review, Actar, FunctionLab, Harvard Graduate School of Design, 2011.
- [6] Ipek Gursel Dino, "Creative design exploration by parametric generative systems in architecture", Journal of the Faculty of Architecture, volume 29, issue 1, 2012.
- [7] John Frazer, "Parametric computation: history and future", Architectural Design, Volume 86, Issue 2, 2016.
- [8] Christiane M. Herr, Thomas Kvan, "Adapting cellular automata to support the architectural design process", Automation in Construction, volume 16, issue 1, 2007.
- [9] Sheng-Fen Chien, "Supporting information navigation in generative design systems", Carnegie Mellon University, 1998.
- [10] Achim Menges, Sean Ahlquist "Computational Design Thinking", Wiley; 1st edition, 2011.
- [11] Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, "Introduction to Algorithms, Second Edition", The MIT Press, 2001.
- [12] Michael J. Ostwald, "Systems and enablers: modeling the impact of contemporary computational methods and technologies on the design process", Computational design methods and technologies, 2012.
- [13] Javier Monedero, "Parametric design: a review and some experiences" Automation in Construction, volume 9, issue 4, 2000.
- [14] Abdelsalam, Mai, "The Use of the Smart Geometry through Various Design Processes: Using the programming platform (parametric features) and generative components", International Conference Proceedings of the Arab Society for Computer Aided Architectural Design ASCAAD, 2009.
- [15] Carlos Roberto Barrios Hernandez, "Thinking parametric design: introducing parametric Gaudi", design studies, volume 27, issue 3, 2006.
- [16] David Karle, Brian M. Kelly, "Parametric thinking", Proceedings of ACADIA regional conference, 2011.
- [17] Tomlow, Jos, "The Model: Antoni Gaudi's hanging model and its reconstruction", PhD thesis, Universität Stuttgart, 1989.
- [18] ÖMER AKIN, "Three Fundamental Tenets for Architectural Ethics, invited paper for the ACSA Teacher's Conference", Cranbrook Academy of Art, 2004.
- [19] John Yeomans, "The Other Taj Mahal: What Happened to the Sydney Opera House", Longman Australia, 1973.
- [20] O ARUP, R S JENKINS, "The evolution and design of the Concourse at the Sydney Opera House" volume 39, issue 4, 1968.

# Calculating the Lower Angular Excited States in Two Dimensions Using the Finite Difference Time Domain Method

Huwaida Elgweri  
Department of Physics  
University of Tripoli  
Tripoli, Libya  
H.Elgweri@uot.edu.ly

Amal Hamed  
Department of Physics  
University of Tripoli  
Tripoli, Libya  
Amal.HAMED@uot.edu.ly

Mohamed Mansor  
Department of Physics  
University of Tripoli  
Tripoli, Libya  
m.mansor@uot.edu.ly

**Abstract** - The Finite Difference Time Domain method has been used to find the angular excited states wave functions in two dimensions. These excited states are calculated by applying the iterative procedure on a specified initial guess wave function that contains the desired excited state as a lowest state, this is simply done by introducing lines of zeros in the wave functions and their second derivatives. This of course depends on the symmetry of the potential. We choose here either square or cylindrical symmetry, so the lowest angular excitations will contain lines of zeros one or two passing through the region, namely, the first excited state and the second excited state respectively. In our investigation, we apply this technique to two simple potentials, which are the two dimensional simple harmonic oscillator and the finite cylindrical well potential in order to illustrate the accuracy and the efficiency of these calculations. These potentials were chosen, as the analytical solutions are available, so to compare them with our results using MATLAB program.

**Keywords:** Finite Difference Time Domain Method, diffusion equation, cylindrical well potential, simple harmonic oscillator, Schrödinger equation

## I. Introduction

The Finite Difference Time Domain Method (FDTD) has several applications such as electromagnetic wave simulations, solving Maxwell's equations [1], solar cells, filters, optical switches, semiconductor based photonic devices and nonlinear devices [2], it is also used for solving Schrodinger equation which is the main topic of our investigation. There are many advantages of this method, for instance; the accuracy of numerical modeling, flexible for any geometrical shapes, and easy in programming [3,4]. However, this method is suitable to calculate the ground state of the quantum systems because its diffusion behavior and in this paper we present a modified technique to improve the (FDTD) to be valid to calculate the lower angular excited states as well.

The time- dependent Schrödinger equation provides a description of the quantum system, that it is given by,

$$i\hbar \frac{\partial \Psi(\vec{r}, t)}{\partial t} = \hat{H} \Psi(\vec{r}, t), \quad (1)$$

where  $\hat{H}$  is the Hamiltonian of the system. The first term of the Hamiltonian contains the kinetic operator and the second

term contains the potential, so the dimensionless form of the Hamiltonian is given by,

$$\hat{H} = -\nabla^2 + V(\vec{r}). \quad (2)$$

The exact eigenfunctions and eigenvalues of the differential equation (1) can be obtained only analytically for a handful of potentials and in most cases we have to resort to the numerical analysis. One of the most useful numerical techniques is the diffusion method, which is based on the equivalence of the time- dependent Schrödinger equation to the diffusion type equation by performing a transformation from real time domain to imaginary time domain  $\tau = \frac{i}{\hbar} t$  to get the following equation

$$\hat{H} \Psi(\vec{r}, \tau) = -\frac{\partial \Psi(\vec{r}, \tau)}{\partial \tau}. \quad (3)$$

There have been various numerical methods to solve this diffusion type equation such as diffusion Monte Carlo method [5], Grimm and Storer approximation method [6], and finite difference time domain method (FDTD) [7]. All these methods involve an iterative procedure applied on an arbitrary initial guess wave function that contains a mixture of all possible state wave functions. This iterative process can be viewed physically as cooling the system and lowering its energy [8], so the iterative procedure will always lead to the ground state of the system. Hence if one only interested in the ground state of a system, the diffusion method is simple and sufficient.

The advantage of this work is to extract higher angular excited states using lines of zeros in the wave function. This is done by classifying the initial guess wave function to even parity wave function and odd parity wave function. This procedure will still give the lowest possible excited state. The space used in these calculations is extended to the unclassical region and it is kept small by using the end point formula for the second derivative. The end point formula allows us to calculate the wave function self consistently in the given region. However, the region should be extended far enough to calculate accurate energy eigenvalue. Due to the symmetry only one sign region is kept for the actual numerical calculations.



In addition to the introduction section this paper is organized into three sections as follows: The general theory section which presents the formulation of the (FDTD) method that is used to calculate the ground state, followed by the improvement performance that is used to employ this method to calculate the lower angular excited states. In application section the improved (FDTD) method is applied to two familiar examples in order to test this technique. Finally, the conclusion section contains the conclusion and the summary in this work.

## II. General Theory

By using separation of variables technique we get the formal solution of the diffusion equation (3) in two dimensions as

$$\Psi(x, y, \tau) = \sum_{i=0}^{\infty} c_i \varphi_i(x, y) \exp(-E_i \tau), \quad (4)$$

where  $c_i$  are expansion coefficients,  $\varphi_i(x, y)$  and  $E_i$  are a complete set of eigenfunctions and their corresponding energy eigenvalues for the time-independent Schrödinger equation, so they satisfy

$$\hat{H} \varphi_i = E_i \varphi_i. \quad (5)$$

In this section, we introduce a short overview of the (FDTD) method to prove that the iterative process leads to the lowest state of the system as follows.

Eq (4) can be rewritten in a convenient form in order to solve it iteratively as,

$$\Psi(x, y, n\Delta\tau) = \sum_{i=0}^{\infty} c_i \varphi_i(x, y) \exp(-E_i n\Delta\tau), \quad (6)$$

where  $\Delta\tau$  is the imaginary time step and  $n$  is an integer presents the number of iteration steps. The iterative procedure is performed using an arbitrary initial guess wave function which is given by,

$$\Psi(x, y, 0) = \sum_{i=0}^{\infty} c_i \varphi_i(x, y). \quad (7)$$

This initial guess wave function contains a mixture of all the wave functions of the system, and the iterative process increases the value of the imaginary time domain in each iteration by value equal  $\Delta\tau$  [9]. Therefore, for a sufficient number of iterative steps which give a large value of  $n\Delta\tau$  that is much larger than the inverse excitation energy only the lowest state wave function remains as a result and the other states damp faster because of their larger eigenvalues, so

$$\lim_{n \rightarrow \infty} \Psi(x, y, n\Delta\tau) \approx c_0 \varphi_0(x, y) \exp(-E_0 n\Delta\tau), \quad (8)$$

where  $E_0$  is the smallest eigenvalue. So, the iterative process is viewed physically as cooling the system and lowering its energy. A normalization of the wave function after each time step ensures the final result being  $\Psi_0(x, y)$ .

Restricting the numerical calculations in the first quarter of  $x - y$  plane allows us to introduce a specific initial guess

wave function that is subjected to certain symmetric properties, and then applying a suitable iterative procedure will converge to the lowest angular excited state in this initial guess wave function.

Therefore, introducing odd initial guess wave function that contains one line of zero lies on either  $x - axis$  or  $y - axis$  will exclude all even state wave functions, so, this odd initial guess wave function will be a mixture of only the odd wave functions of the system as

$$\Psi_{od}(x, y, 0) = \sum_{i=0}^{\infty} c_{2i+1} \varphi_{2i+1}(x, y), \quad (9)$$

since only the lowest state remains after applying the iteration, it can easily be seen that applying iterative procedure which subjected to the anti symmetric property on the zero line will approach to the first angular excited state i.e.

$$\lim_{n \rightarrow \infty} \Psi_{od}(x, y, n\Delta\tau) \approx c_1 \varphi_1(x, y) \exp(-E_1 n\Delta\tau). \quad (10)$$

Similarly, introducing even initial guess wave function that contains two lines of zeros lay on both  $x - axis$  and  $y - axis$  will exclude all dissimilar state wave functions as

$$\Psi_{ev}(x, y, 0) = \sum_{i=0}^{\infty} c_{2i+2} \varphi_{2i+2}(x, y), \quad (11)$$

and applying appropriate iteration procedure will approach to the second angular excited state i.e.

$$\lim_{n \rightarrow \infty} \Psi_{ev}(x, y, n\Delta\tau) \approx c_2 \varphi_2(x, y) \exp(-E_2 n\Delta\tau). \quad (12)$$

As stated before, the norm of the total wave function is kept equal to one before the next iteration step in order to prevent the wave function from diminishing, since the iterative process damps all wave functions but with different rates.

Detailed finite difference scheme simulation in two dimensions based on symmetric and anti-symmetric properties will be presented hereafter.

Following the forward finite difference formula, the first order time derivative in (3) can be approximately written as

$$\frac{\partial \Psi(x, y, \tau)}{\partial \tau} = \frac{\Psi(x, y, \tau + \Delta\tau) - \Psi(x, y, \tau)}{\Delta\tau}. \quad (13)$$

Similarly, the Laplacian operator in (2) can be found by the central finite difference formula as

$$\begin{aligned} \nabla^2 \Psi(x, y, \tau) = & \frac{\Psi(x + \Delta x, y, \tau) + \Psi(x - \Delta x, y, \tau) - 2\Psi(x, y, \tau)}{\Delta x^2} \\ & + \frac{\Psi(x, y + \Delta y, \tau) + \Psi(x, y - \Delta y, \tau) - 2\Psi(x, y, \tau)}{\Delta y^2}, \end{aligned} \quad (14)$$

where  $\Delta x$  is the mesh size between adjacent points  $\Delta x = x_{n+1} - x_n$ , and  $\Delta y$  is the mesh size between adjacent points  $\Delta y = y_{n+1} - y_n$ . While the Laplacian operator at the axis depends on the symmetric property along this axis. If the axis includes zero line then the second derivative along this axis equal to zero i.e.

$$\nabla^2 \Psi(x, 0, \tau) = 0, \quad (15)$$

$$\nabla^2 \Psi(0, y, \tau) = 0. \quad (16)$$

If there is no zero line on the axis then the second derivative across this axis is given by

$$\nabla^2 \Psi(x, 0, \tau) = \frac{\Psi(x + \Delta x, 0, \tau) + \Psi(x - \Delta x, 0, \tau) - 2\Psi(x, 0, \tau)}{\Delta x^2} + \frac{2\Psi(x, 1, \tau) - 2\Psi(x, 0, \tau)}{\Delta y^2}, \quad (17)$$

$$\nabla^2 \Psi(0, y, \tau) = \frac{2\Psi(1, y, \tau) - 2\Psi(0, y, \tau)}{\Delta x^2} + \frac{\Psi(0, y + \Delta y, \tau) + \Psi(0, y - \Delta y, \tau) - 2\Psi(0, y, \tau)}{\Delta y^2}. \quad (18)$$

The second order spatial derivative at the boundaries of the spatial mesh is calculated using the end point difference formula as

$$\frac{\partial^2 \Psi(x_N, y, \tau)}{\partial x^2} = \frac{2\Psi(x_N, y, \tau) - 5\Psi(x_{N-1}, y, \tau) + 4\Psi(x_{N-2}, y, \tau) - \Psi(x_{N-3}, y, \tau)}{\Delta x^2}, \quad (19)$$

and

$$\frac{\partial^2 \Psi(x, y_N, \tau)}{\partial y^2} = \frac{2\Psi(x, y_N, \tau) - 5\Psi(x, y_{N-1}, \tau) + 4\Psi(x, y_{N-2}, \tau) - \Psi(x, y_{N-3}, \tau)}{\Delta y^2}. \quad (20)$$

The potential operator in (2) can be written as

$$V(x, y) \Psi(x, y, \tau) = V(x, y) \frac{\Psi(x, y, \tau) + \Psi(x, y, \tau + \Delta \tau)}{2}. \quad (21)$$

By plugging (13), (14) and (21) into (3) we get the recursion form equation

$$\begin{aligned} \Psi^{n+1}(i, j) &= \alpha(i, j) \Psi^n(i, j) \\ &+ \beta(i, j) \frac{\Delta \tau}{\Delta x^2} [\Psi^n(i+1, j) + \Psi^n(i-1, j) - 2\Psi^n(i, j)] \\ &+ \beta(i, j) \frac{\Delta \tau}{\Delta y^2} [\Psi^n(i, j+1) + \Psi^n(i, j-1) - 2\Psi^n(i, j)], \end{aligned} \quad (22)$$

where the notation  $\Psi^n(i, j) \equiv \Psi(i\Delta x, j\Delta y, n\Delta \tau)$ ,  $\alpha(i, j)$  and  $\beta(i, j)$  are given respectively by

$$\alpha(i, j) = \frac{1 - \Delta \tau V(i, j)}{1 + \Delta \tau V(i, j)}, \quad (23)$$

$$\beta(i, j) = \frac{1}{1 + \Delta \tau V(i, j)}. \quad (24)$$

For the purpose of numerical computation, the recursions equation (22) must converge at the chosen imaginary time steps, therefore the coefficients  $\frac{\Delta \tau}{\Delta x^2}$  and  $\frac{\Delta \tau}{\Delta y^2}$  must satisfy the stability condition which is given in two dimensions by [6]

$$\frac{\Delta \tau}{\Delta x^2} + \frac{\Delta \tau}{\Delta y^2} \leq 1. \quad (25)$$

In other words that  $\Delta \tau$  must be chosen carefully because of its association with the grid spacing.

The energy eigenvalues are calculated by means of numerical evaluation of the expectation value of the Hamiltonian for their corresponding normalized eigenfunctions as

$$E = 4 \sum_i^N \sum_j^N \Psi^*(i, j) \hat{H} \Psi(i, j) \Delta x \Delta y. \quad (26)$$

### III. Applications

#### A. Two Dimensional Simple Harmonic Oscillator

As a first example, we consider the simple harmonic oscillator in two dimensions which is a good example to test the validity of the presented method. The distance is measured in unit  $\sqrt{\frac{\hbar}{m\omega}}$  and the unit of the energy is  $\frac{1}{2}\hbar\omega$ . In these units the Hamiltonian operator for the relative system is

$$\hat{H} = -\nabla^2 + r^2, \quad (27)$$

where  $\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$ , and  $r^2 = x^2 + y^2$ ,

the exact eigenfunctions are

$$\Psi_{nx, ny}(x, y) = \frac{1}{\sqrt{nx! ny! 2^{(nx+ny)} \pi}} \left[ x - \frac{d}{dx} \right]^{nx} \exp\left(-\frac{1}{2}x^2\right) \left[ y - \frac{d}{dy} \right]^{ny} \exp\left(-\frac{1}{2}y^2\right), \quad (28)$$

and the exact energy eigenvalues are

$$E(nx, ny) = 2(nx + ny + 1). \quad (29)$$

The first three lowest eigenfunctions of the simple harmonic oscillator calculated numerically are shown in Figure I (a,b,c), respectively. The figures have been provided with a comparison of the numerical wave functions with their corresponding exact wave functions to confirm the accuracy of the numerical method that we have adapted. The generated first angular excited state presented in Figure I (b) is  $\Psi_{1,0}(x, y)$ , and the second angular excited state presented in Figure I (c) is  $\Psi_{1,1}(x, y)$ . The associated comparison figures show the difference of the numerical wave functions from their corresponding exact wave functions, these comparison figures illustrate that the maximum value of the error is nearly about  $\mp 10^{-4}$  which reflects the high accuracy of this method. Noteworthy that applying symmetric technique produces the parts of eigenfunctions in the first quarter in  $x - y$  plane, therefore, the numerical cost will be reduced to the quarter of its value required to obtain the eigenfunctions in whole  $x - y$  plane. The presented contour plots of the eigenfunctions are calculated by extending the eigenfunctions that are obtained using symmetric technique over the whole  $x - y$  plane.

In Table I, we show the numerical energy eigenvalues for the first three lowest states obtained using (26) and their corresponding exact energy eigenvalues which they are in a good agreement.

Table I. Comparison of the numerical energy eigenvalues of the first three lowest states with their corresponding exact energy eigenvalues for two dimensional simple harmonic oscillator

The state	Numerical eigenvalue	Exact eigenvalue
Ground state	1.9999	2.0
First excited state	3.9977	4.0
Second excited state	5.9953	6.0

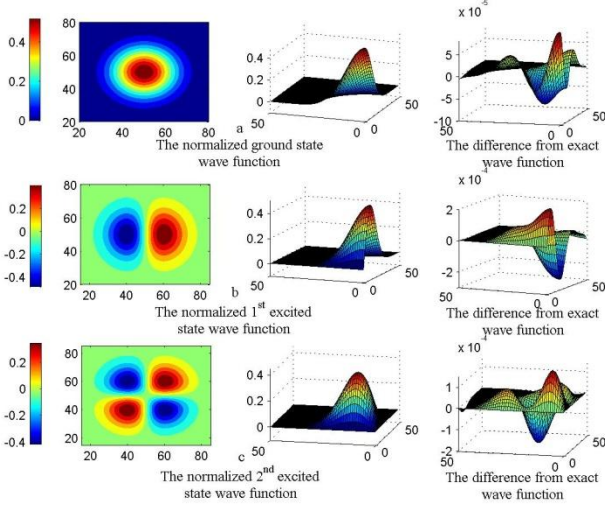


Figure I. The first three lowest states wave functions of two dimensional simple harmonic oscillators calculated numerically.

- The normalized ground state wave function  $\Psi_{0,0}(x, y)$ .
- The normalized first angular excited state wave function  $\Psi_{1,0}(x, y)$ .
- The normalized second angular excited state wave function  $\Psi_{1,1}(x, y)$ .

All the previous calculations are performed with  $\Delta\tau = 0.001$  and  $\Delta x = \Delta y = 0.1$ , in these parameters 2000 iterations are sufficient to get acceptable results. Numerically, the integrals calculations used to normalize the wave function and those used to determine the energy eigenvalues are evaluated using the trapezoidal rule.

### B. Finite Cylindrical Well Potential.

As a second example we extend the same calculations to the finite cylindrical well potential which is given in Cartesian coordinate by

$$V(x, y) = \begin{cases} -V_0 & x^2 + y^2 \leq r_0^2 \\ 0 & x^2 + y^2 > r_0^2 \end{cases} \quad (30)$$

where  $V_0$  is the depth of the potential, and  $r_0$  is the radius of the potential.

In the distance unit  $a$  and energy unit  $\frac{\hbar^2}{2ma^2}$  the dimensionless form of the time independent Schrödinger equation is given by,

$$-\nabla^2 \Psi(x, y) + v(x, y) \Psi(x, y) = \varepsilon \Psi(x, y), \quad (31)$$

where  $v(x, y) = \frac{2ma^2}{\hbar^2} V(x, y)$ , and  $\varepsilon = \frac{2ma^2}{\hbar^2} |E|$ .

The exact energy eigenvalues and their corresponding eigenfunctions are calculated analytically by transforming (31) into the polar coordinate formula as

$$-\frac{1}{r} \left( \frac{\partial}{\partial r} r \frac{\partial}{\partial r} \right) \Psi(r, \theta) - \frac{1}{r^2} \frac{\partial^2 \Psi(r, \theta)}{\partial \theta^2} + v(r, \theta) \Psi(r, \theta) = \varepsilon \Psi(r, \theta). \quad (32)$$

By using separation of variables technique we get the solutions of the second order partial differential equation for cylindrical potential as [10]

$$\Psi(r, \theta) = \begin{cases} J_m(\sqrt{v_0 - \varepsilon} r) \exp(\pm im\theta) & r \leq r_0 \\ K_m(\sqrt{\varepsilon} r) \exp(\pm im\theta) & r > r_0 \end{cases}, \quad (33)$$

where  $J_m(\sqrt{v_0 - \varepsilon} r)$  is the Bessel function of the order  $m$ ,  $K_m(\sqrt{\varepsilon} r)$  is the modified Bessel function of the order  $m$ , and  $m$  is an integer which for  $m = 0$  corresponds to the ground state wave function,  $m = 1$  corresponds to the first angular excited state wave function and  $m = 2$  corresponds to the second angular excited state wave function. In order to obtain the energy eigenvalues we apply the continues conditions at the well boundary to get the following equation,

$$\sqrt{v_0 - \varepsilon} J'_m(\sqrt{v_0 - \varepsilon} r_0) K_m(\sqrt{\varepsilon} r_0) - \sqrt{\varepsilon} K'_m(\sqrt{\varepsilon} r_0) J_m(\sqrt{v_0 - \varepsilon} r_0) = 0, \quad (34)$$

where  $J'_m(s)$  defined as  $\frac{dJ_m(s)}{ds}$ , and  $K'_m(s)$  defined as  $\frac{dK_m(s)}{ds}$ .

Therefore, the algebraic equation (34) will be satisfied only at certain discrete energy eigenvalues  $\varepsilon$ , which can be obtained by getting the roots of this equation numerically. After energy eigenvalues have been found their corresponding eigenfunctions can be calculated very easily by plugging the energy eigenvalue into (33).

Numerically, in the finite cylindrical well potential we have to use small spatial mesh size because of the circularly shape of this potential, thereby the time step must be chosen carefully depending on the stability condition. In Tables (II, III and IV), we show the effect of reducing the values of spatial mesh size and the associated time step on the numerical eigenvalues, by presenting the first three lowest states of finite cylindrical well potential with depth  $v_0 = 30$ , and radius  $r_0 = 2$ . However, smaller time step requires larger number of iterations to get acceptable results.

In Figure II (a,b,c), we show the first three lowest eigenfunctions respectively for finite cylindrical well potential with depth  $v_0 = 30$  and radius  $r_0 = 2$ , these eigenfunctions are calculated numerically using the most fit parameters. To illustrate the accuracy of the numerical results we show the difference between the numerical eigenfunctions and their corresponding exact eigenfunctions in front of the figures. In this case to get acceptable results nearly 5000 iterations are required. Again, the integrals are evaluated numerically using the trapezoidal rule.

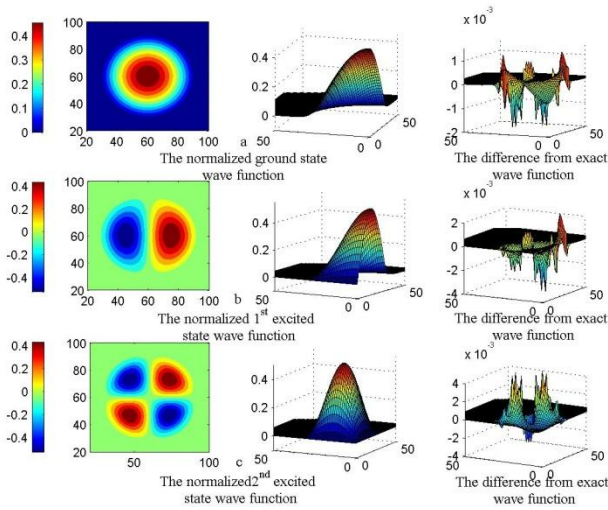


Figure II. The first three lowest states wave functions of the finite cylindrical well calculated numerically.  
a. The normalized ground state wave function  $\Psi_{0,0}(x, y)$ .  
b. The normalized first angular excited state wave function  $\Psi_{1,0}(x, y)$ .  
c. The normalized second angular excited state wave function  $\Psi_{1,1}(x, y)$ .

Table II. The numerical ground state eigenvalue of finite cylindrical well potential with  $v_0 = 30$  and  $r_0 = 2$  calculated using different values of  $\Delta x, \Delta y$  and  $\Delta \tau$ . The analytical eigenvalue is  $-28.78743$

$\Delta x = \Delta y$	$\Delta \tau$	Numerical eigenvalue	Absolute error
0.1	0.001	-28.79096	0.00353
0.08	0.0005	-28.78791	0.00048
0.07	0.0004	-28.78778	0.00035

Table III. The numerical first angular excited state eigenvalue of finite cylindrical well potential with  $v_0 = 30$  and  $r_0 = 2$  calculated using different values of  $\Delta x, \Delta y$  and  $\Delta \tau$ . The analytical eigenvalue is  $-26.92663$

$\Delta x = \Delta y$	$\Delta \tau$	Numerical eigenvalue	Absolute error
0.1	0.001	-26.93885	0.01222
0.08	0.0005	-26.92997	0.00334
0.07	0.0004	-26.92901	0.00238
0.05	0.0003	-26.92866	0.00203
0.04	0.0001	-26.92516	0.00147

Table IV. The numerical second angular excited state eigenvalue of finite cylindrical well potential with  $v_0 = 30$  and  $r_0 = 2$  calculated using different values of  $\Delta x, \Delta y$  and  $\Delta \tau$ . The analytical eigenvalue is  $-24.48947$

$\Delta x = \Delta y$	$\Delta \tau$	Numerical eigenvalue	Absolute error
0.1	0.001	-24.55968	0.07021
0.08	0.0005	-24.52546	0.03599
0.07	0.0004	-24.48951	0.00004

## IV. Conclusion

This present investigation has demonstrated the usefulness of the (FDTD) method with the appropriate symmetric boundary conditions of the wave functions to extract the eigenfunctions and eigenvalues of the lower angular excited states for cylindrical potential if they exist. The diffusion method is very suitable for obtaining the ground state calculations and it was proved in this paper that this method is valid for determining the lower angular excited state as well, by choosing an appropriate initial guess wave function that has to be odd function for the first excited state and even function for the second excited state. Choosing special initial guess wave function removes all dissimilar excited states and forced the iterative procedure to yield to the lowest state in this initial guess function. Numerically removing the ground state by Gram Schmidt orthogonalization procedure can also lead to the excited states but the method introduced in this paper offer the advantage over the previous method that it's less expensive in numerical cost. In addition, since the calculations of symmetric method performed in the first quarter of  $x - y$  plane, so it reduces the greatly numerical cost to the quarter of those required to perform the calculations to entire the plane, end point derivatives farther reduce the mesh points considerably. We have provided detailed calculations to illustrate this improved technique in two dimensions, and we applied it to two different potentials as examples, namely the simple harmonic oscillator and the finite cylindrical well. The numerical results illustrated in Figure I, and in Table I, which correspond to the simple harmonic oscillator potential show the efficiency and simplicity of this method, while the numerical results illustrated in Figure II, and in Tables (II, III, and IV), those correspond to the finite cylindrical well potential show that extra care should be taken when choosing the parameters used to obtain the eigenfunctions and the energy eigenvalues in this case. Finally, we can generalize this method to calculate more angular excited states by using diagonal zeros axis.

## References

- [1] Antonio Soriano, Enrique Navarro, Jorge Porti, Vicente Such. "Analysis of the finite difference time domain technique to solve the Schrödinger equation for quantum devices", Journal of Applied Physics, 95 (12), (2004).
- [2] Charles Reinke, Aliakbar Jafarpour, Babak Momeni, Mohammad Soltani, Sina Khorasani, Ali Adibi, Yong Xu, and Reginald Lee, "Nonlinear Finite-Difference Time-Domain Method for the Simulation of Anisotropic,  $\chi(2)$ , and  $\chi(3)$  Optical Effects", Journal of lightwave technology, 24, no. 1, (2006)
- [3] Dennis Sullivan, David Citrin, "Time-domain simulation of two electrons in a quantum dot", Journal of Applied Physics, 89,3841,(2001).
- [4] I Wayan. Sudiarta, Lily. Maysarr. Angraini "The Finite Difference Time Domain (FDTD) Method to Determine Energies and Wave Functions of Two-Electron Quantum Dot" in Proc. AIP LLC, 2023, p. 020199-1, 2018.
- [5] Thiago N Barbosa, Marcos M Almeida, Frederico V Prudente, "A quantum Monte Carlo study of confined quantum systems: application to harmonic oscillator and hydrogenic-like atoms", Journal of Physics B: Atomic, Molecular and Optical Physics, 48 (5), (2015).

- [6] R.Grimm, R.G. Storer, "A new method for the numerical solution of the Schrödinger equation", *Journal of Computational Physics*, 4, 230-249, (1969).
- [7] I Wayan Sudiarta, D.J.Wallace Geldart, "Solving the Schrödinger equation using the finite difference time domain method", *Journal of Physics A: Mathematical and Theoretical*, 1885-1896, (2007).
- [8] Huwaida Elgweri, Mohamed Mansor, "First excited solutions of Schrödinger equation by the diffusion method applied to various one dimension problem", *Journal of Academy for Basic and Applied Science*, 14 (1), 1-4, (2015).
- [9] Mohamed Mansor, Taher Sherif, Saleh Swedan, "Improved simple numerical method using the diffusion equation applied for central force bound quantum systems", *Journal of basic and applied science*. 14,72-81,(2004).
- [10] George Arfken, Hans Weber, *Mathematical Methods for Physics*. 7<sup>th</sup> edition. Academic Press is an imprint of Elsevier.225 Wyman Street, Waltham, MA 02451, USA.
- [11] Huwaida Elgweri, Mohamed Mansor,"Calculation of positive spectrum for the higher excited states using Grimm and Storer diffusion method", *The Libyan journal of science*. 30, 33-42, (2017).
- [12] Kailash Kumar, "On expanding the exponential". *Journal of Mathematical Physics*. 6, 1928-34, (1965).
- [13] Ivan Sokolnikoff, Raymond Redheffer, *Mathematics of Physics and Modern Engineering*, 2<sup>nd</sup> edition, McGraw-Hill, New York, (1966).
- [14] Robert Eisberg, Robert Resnick, *Quantum Physics of Atoms, Molecules, Solids, Nuclei, and particles*, John Wiley and Sons, New York, (1974).

# Implementation Framework for a Blockchain-Based Reputation and Trust System

Zeba Mahmood

Department of Software Engineering  
Kaunas University of Technology  
Zeba.mahmood@ktu.edu

Vacius Jusas

Department of software Engineering  
Kaunas University of Technology  
Vacius.Jusas@ktu.lt

**Abstract**— *The web revolutionized how people engage with data. Consumers in e-commerce rely solely on online reputation systems when selecting which items to purchase. For its consumers and sellers, many e-commerce platforms now include built-in review systems. Users may share their evaluations and even propose better items for their online shopping on social media networks like Facebook. Despite these advancements, there are still significant flaws numerous platforms that facilitate online interactions are still centralized and vulnerable to manipulation, tends to result in a broken marketplace with ineffective verification where vendors can easily manipulate consumers' perceptions to increase sales. Blockchain has been hailed as a game-changing technology that may bring an extra layer of trust and security to online interactions, as well as providing much-needed reputation in online interaction platforms. We offer a trustworthy, decentralized reputation model based on the Ethereum Blockchain in this paper, with the goal of restoring trust and integrity in the online interaction industry. We go over the implementation framework for such a system and present some preliminary results. Our findings indicate that a decentralized Blockchain-based reputation network is feasible, with impact factor evaluations for each node serving as the primary criterion for assessing the ecosystem's trustworthiness.*

**Keywords**— *Reputation, Blockchain, trust, peer-to-peer, consortium networks, smart contracts*

## I. INTRODUCTION

The internet and social media have been acknowledged as new frontiers for media convergence a phenomenon that is increasingly characterized by how information flows and users migrate linking content, communication, and computation in a complex setup. Users rely on the Internet to send and receive emails, search online for media files or news, and shop for products and services. As of this writing, the number of internet users has surged to 4.4 billion from 4.3 billion in 2018 [1],[2]. At this rate, online interactions occur between known entities. However, this is not the case. Online communication takes place in an environment where entities are anonymous with no mechanisms for verifying interactions between them. Entities can easily create many accounts for each platform and engage in online interactions. In e-commerce and social media platforms, this has led to the emergence of a thriving economy based on fake reviews. Driven by profits, many merchants and vendors are buying positive reviews for their businesses and negative reviews for their rivals in an attempt to influence how users perceive their businesses [3],[4]. In 2017, the U.K. Consumer Advocacy Group reported that sellers on Amazon were listing products and services that carried thousands of positive fake reviews [5]. This problem is prevalent because many of these platforms are centralized and prone to manipulation. As such, the online marketplace is broken and

lacks a robust verification system allowing vendors to easily manipulate reviews to influence consumers' perception about their products. Major contributions of this study are three-fold.

- Contextualize reputation and trust systems regarding Blockchain technology.
- Propose a model consisting of actors in the context of reputation systems, the network architecture, and components.
- Propose an Ethereum-based platform that implements the reputation model.

The rest of paper is organized as follows: section II Blockchain technology discusses the reputation model and provides an implementation framework while section III discusses the initial practical findings.

### A. ONLINE INTERACTION MARKETPLACE IN THE CONTEXT

Online interaction marketplace is currently broken because many reputation systems that are in use are centralized. As such, there are no mechanisms for guaranteeing that the behavior of entities remains honest during the interaction process. In recent times, calls for a decentralized Internet have been growing. Although the modern Internet is built on top of decentralized protocols such as TCP/IP and HTTP, a large section of the application stack has remained centralized. The desire for more decentralization has largely emanated from the broken marketplace bedeviled by fraudulent activities. Blockchain can help enforce reputation and trust in online interactions. Such a platform can help predict the outcomes of online interactions. At present, there are many online businesses that have successfully implemented decentralized computing systems with disruptive consequences. The unveiling of Bitcoin in 2009 led to the emergence of decentralized alternative platforms like Open Bazaar and Silk Road. The success of these platforms has been appraised by their ability to create a trustless economy that does not require trust third parties for transaction verifications. We believe that Blockchain holds the potential to unlock the problems bedeviling the reputation systems. Our model provides a framework for implementing an open and transparent reputation system based on Ethereum Blockchain. With this system, any user intending to participate in online interaction with the other party can verify the true identities and proof that the parties are who they claim to be.

### B. Blockchain Technology

Blockchain was first introduced with the publication of whitepaper in 2008 by Satoshi Nakamoto as break-through.

technology for Bitcoin the first peer-to-peer (P2P) cryptocurrency [6]. The technology paired cryptography an already established and understood concept in computer science at the time with a novel Merkle tree data structure to facilitate digital transactions. What made the technology grow fast is its ability to solve the decade-old double-spending problem: a scenario where the same money is copied and spent more than once. With a P2P model and Merkle tree data structure, Blockchain does not require intermediaries, such as banks, to facilitate online transactions. Blockchain became the progenitor of cryptocurrencies with Bitcoin becoming its first use case. Rather than having traditional accounting systems of banks and other intermediaries to validate online transactions, cryptocurrencies use a public digital ledger or register (Blockchain) to confirm transactions. When users transact, the transaction is recorded on the Blockchain. With traditional accounting systems, security and validation of the register depend on banks, central banks, card issuers, and lately telecommunication firms for mobile payments. However, in Blockchain-backed transactions, the same is handled by decentralized nodes (also called miners) who compete for verification of the transactions by solving a complex mathematical puzzle. In essence, what Blockchain managed to solve is replacing trusted centralized authority's entities with the decentralized and trustless system.

Inspired by Bitcoin, other cryptocurrencies (generally referred to as altcoins) emerged with Ethereum becoming the dominant platform. Ethereum was unveiled in 2013 not just as a digital cash system but rather as a programmable platform with unconceivable capabilities. With Ethereum, smart contracts and other Decentralized applications (DApps) could be executed in a "complete Turing World Computer." Ideally, smart contracts can be regarded as a special type of accounts that are recorded on the Blockchain and are therefore not controlled by humans. In their most basic forms, smart contracts can run all kinds of instructions like maintaining states, checking conditions, and sending and receiving digital money. Of utmost significance is that the smart contract on the Blockchain cannot be changed, and/or even hacked. These attributes make Blockchain such as Ethereum a perfect platform for enforcing reputation rules because it is a permission-less Blockchain providing designated members with the ability to read and write on the ledger [7].

### C. Consensus Mechanisms

In Blockchain, consensus algorithms form the basic rules of agreement on how the nodes in the network validate the transactions [8]. For example, suppose Alice sends \$10 worth of bitcoins to Bob. There has to be a mechanism in place that ensures that Alice's account balance reduces by \$10 while Bob's account balance increases by the same amount. Such a mechanism has to be implemented in a manner that does not allow any malicious transactions or alterations to the Blockchain to be made without the full consent of all the nodes participating in the network. Some of the most common consensus algorithms include.

#### a) PoW (Proof-of-Work)

It is by far the most used consensus mechanism. It was first applied in Bitcoin. In a PoW, miners compete against each other by solving a complex.

mathematical puzzle to authenticate and transactions and append the new blocks to the chain and in the process by they are rewarded by coins. Whereas computation processes leading to new blocks are difficult, they can easily be verified by nodes in the Blockchain ecosystem. As such, when a miner obtains the solution to a new block, it broadcasts the generated block to the network for verification by other miners [9]. All the other miners need to confirm that the solution is correct for the generated block to be confirmed. The mathematical puzzle to be solved in PoW include [10].

- Generating a hash function which requires the miner to determine input if the output is known.
- Integer factorization which involves presenting a number as a product of two other integers (usually large primes).
- Checking to confirm whether a DoS attack has occurred by computing hash functions.

A node that successfully generates a block is rewarded with coins as a form of incentive. PoW helps to protect the Blockchain network against attacks since an attack can only occur with a lot of computational power and time which would be inefficient since it would be expensive in terms of costs than the potential rewards.

#### b) PoS (Proof-of-Stake)

In a PoS,[11] the consensus is achieved by requiring nodes to stake some of their coins or tokens in the process of authenticating blocks. Essentially, staking involves depositing some coins which are then locked up into the Blockchain ecosystem. Such coins become collateral for vouching for the new block [12]. The more a particular nodes stakes in the ecosystem, the better the chances of being selected to validate the transactions [13]. PoS is specially conceived to resolve the Byzantine Fault Tolerance issue that is rampant with PoW algorithm since all the validators are known in the network and can easily be tracked on the Blockchain [14].

### D. Types of Blockchains

Blockchains are broadly grouped into 3: Public permission-less Blockchains; Public permissioned Blockchains; and Private permissioned Blockchains [15]. In a Public permission-less Blockchain, there is no centralized entity that authorizes the transactions on the Blockchain. These Blockchains can be regarded as shared public ledgers where any node can view and modify the data so long as the node is participating in the network. Ethereum and Bitcoin are some of the known examples under this category. In Public permissioned Blockchain, selected nodes are used to authenticate the transactions on the Blockchain. For example, authentication of the transactions can be assigned to a government entity, senior employees, or an institution. Lastly, a Private Permissioned Blockchain is a ledger ecosystem where data is not available for public view.

### E. Smart contracts

Smart contracts are essentially special types of accounts that are recorded on the Blockchain and are therefore not controlled by humans [15]. In their most basic form, smart contracts can run all kinds of instructions like maintaining

states, checking conditions, and sending and receiving digital money. Of utmost significance is that the smart contract on the Blockchain cannot be changed, and/or even hacked [16]. Ethereum is one example that was developed to help developers to program smart contracts besides acting as a cryptocurrency. With Ethereum, smart contracts and other Decentralized applications can be executed in a “complete Turing World Computer [17]. A permission-less Ethereum Blockchain can help enforce reputation rules since they can control who can read and write on the ledger regarding data being managed [18], [19].

### F. Blockchain and Reputation systems

A reputation system that is open and transparent should be able to compute the trustworthiness of an entity in an online interaction process. A decentralized reputation system incorporating a smart contract provides standardized mechanisms for accessing reputation data which has been aggregated where more authentications reinforce the notion of reliability for anyone reputation of the score. Meanwhile, Blockchain’s consensus mechanisms in the reputation platform can also help to safeguard against known attacks in centralized systems such as Sybil attacks, whitewashing, Denial of service attacks, and slandering [20]. Transacting parties can leverage reputation indicators to decide who to interact with and who to avoid. In e-commerce, merchants and vendors are incentivized to sell credible products and services [21].

## II. IMPLEMENTATION FRAMEWORK

### A. REPUTATION MODEL

A model that captures the reputation of an entity in an online interaction should consists of both endorser and endorsee interacting as shown below:

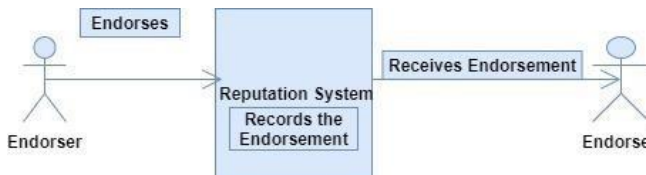


Figure 1. Reputation Model for online Interaction.

The acquaintance process would work as follows:

- Entities A and B have personally known each for a long time. They probably have worked together, or went to the same school, and therefore, are acquaintances of each other.
- Entity A may have interacted so many times with entity B. Perhaps this interaction was in online shopping where A and B successfully transacted and, in the process, A establishes that B is a credible seller while B establishes that A is the credible buyer.

Based on past experiences that A had with B, A is more likely to endorse B on the reputation system. If A endorses B, then this means that A has built trust in B. In this case, the endorsement acts as a transaction message originating from a user account on a Blockchain and is destined to another user account (endorsee). In the network, every user manages two distinct lists of entities that they have so far interacted with: list of endorsers and list of endorsees. The list must have account addresses that identify that user on the platform. The reputation platform would record all the endorsements between the endorsers and endorsees. The system then aggregates the information and compute a total trust score.

The total trust score represents the impact that a particular entity has on the reputation platform. Each entity has two assignment variables:

- $E_g$ : Total endorsements given by the entity; and
- $E_r$ : Total endorsement received by the entity.

To be considered as a node in the network, each entity must have  $E_g$  and  $E_r$  values equal to 1. The methodology used to model the reputation is described as follows:

- **Re**: Ratio of  $E_g$  to  $E_r$ . It is an indicator of how the total sent and received endorsements are far from one another.  $R_e$  must be less than or equal to 1 and is computed as follows:

$$R_e = \min(E_g, E_r) / \max(E_g, E_r) \dots \dots \dots (1)$$

In a reputation system, the ratio between incoming and outgoing endorsement connections should be maintained. This ratio helps to build a trustworthy behavior where a high ratio value is an indicator that a node participating in the network is a high impact performing and trusted node.

- **CPTs**: Total consumable points for A. Every entity that joins the reputation platform receives an equal CPTs from the platform and the value keeps on depleting with each endorsement and is computed as:

$$CPTs = 1/E_r \dots \dots \dots (2)$$

The reputation system measures the contribution of each node in the network and total consumable points help to provide an indication of this measure. For example, the total consumable points for a particular node can limit that node’s ability to convince endorsees that it is a trustworthy node.

- **RPTs**: Total received points for A. It is the aggregated sum of all the consumable points that are received by A from the endorsers. If endorsers E is the set  $\{e_1, e_2, e_3 \dots e_n\}$  for A of the size of E is n, then RPTs is computed as follows:

$$RPTs = \sum(CPTs) \dots \dots \dots (3)$$

- **IF**: Impact Factor: Indicates the reputation score for A and is computed as:

$$IF = R_e * RPTs \dots \dots \dots (4)$$

Ethereum Blockchain is an ideal platform to be used as an implementation platform because it is public and permission less. As such, an endorser (any entity in an online interaction marketplace) to endorse the endorsee in a public and permission less platform. In 2016, Vitalik Buterin—a Russian mathematician and cryptographer—unveiled Ethereum as an evolving platform to rival the existing Bitcoin platform. At the time, the Bitcoin protocol could only validate the ownership and transfer of coins. Ethereum protocol was revolutionary in the sense that instead of just being applied to enforce smart contracts so long as they have enough Ether (ETH) in their wallets. On Ethereum platform, ETH acts as a native cryptocurrency that performs many Ethereum-based applications.

When a smart contract executes on the Blockchain, all the nodes participating in the Ethereum (miners) execute the same code and needs to be agree on a consensus mechanism [22]. At the time, the Ethereum was based on a consensus mechanism called Proof-of-Work (PoW). In this mechanism, nodes to compete to solve a complex mathematical and cryptographic puzzle based on General Byzantine’s problem [23].



PoW served 2 major purposes: verification of the legitimacy of transactions to avoid the double-spending problem that previously existed with digital currencies; and create new coins where miners who successfully perform computations are rewarded with ETH. However, this approach consumed so much hardware resources and Ethereum has now migrated to a new consensus algorithm called Proof-of-Stake (PoS). In the PoS algorithm, validators (instead of miners) are used to lock up some ETH that acts as a stake in the Ethereum ecosystem [24]. Validators that bet their ETH on the blocks are rewarded with coins that are proportional to their stake in the ecosystem [25].

With its smart contract's potential, Ethereum has now become a massive, decentralized platform—what is often called a “complete Turing machine” or Ethereum Virtual Machine (EVM). Ethereum is the dominant platform that supports both public/private management of transactions, miners are required to deposit some currency (Ether) to validate the transaction in an arrangement known as Proof-of-Stake (PoS) [26]. An EVM is Ethereum's interpreter environment that converts smart contracts from high-level language statements into machine language. EVM acts an interpreter for the Ethereum's assembly language. As it is the case with programming in assembly languages, writing codes may be challenging for new developers. Therefore, the Ethereum Foundation proposed Solidity—a high-level language—to be used as the basis for coding smart contracts. Essentially, an Ethereum ecosystem has non-exhaustive elements that are complete with cryptographic tokens, the address of nodes, consensus algorithms (PoS), validators, the Blockchain/Ledger, EVM, and scripting languages are [28], [29].

### B. Reputation Network Architecture

The diagram below summarizes the architecture of the Reputation system:

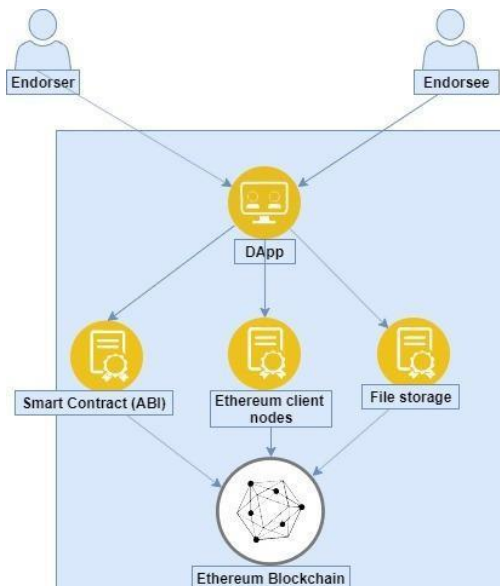


Figure 2. Architecture of Reputation System

The reputation system as presented in this figure 1 having following components.

- DApps;
- Smart contracts; and
- Ethereum Blockchain.

The participants in this reputation system are endorser and endorsee. For example, if Alice has interacted with Bob so many times and wishes to recommend him based on interactions, Alice becomes the endorser while Bob is the endorsee. The endorser uses a DApp (a client web application) which runs a browser to endorse the endorsee in the system. Since these recommendations must be stored on the Blockchain, a smart contract is required. There are three smart contracts that must be generated for the system to work seamlessly Application Binary Interface which is a compiled byte code representation of reputation system, Ethereum client nodes, to manage the nodes joining the network, and file storage that manages the storage of files on the Blockchain. Finally, the system requires a Blockchain which in this case is the Ethereum Blockchain.

#### A. Smart contracts

Smart contracts allow the reputation platform to record all the endorsements between the endorsers and endorseees. They then aggregate the information and compute a total trust score. The total trust score represents the impact factor that a particular entity has on the reputation platform. When smart contracts compile successfully, they generate ABI (Application Binary Interface) which is a binary representation of the compiled EVM. ABI is then deployed to the Ethereum network resulting in the contract obtaining an address and bytecode recorded. The smart contracts are then invoked using Web3.js - a JavaScript API that allows DApps to interact with remote or local Ethereum nodes [30]. The main contract in this system is Endorse contract that defines the logic necessary for any endorsement in the Ethereum network. The Endorse contract can be simplified using the flow chart diagram below:

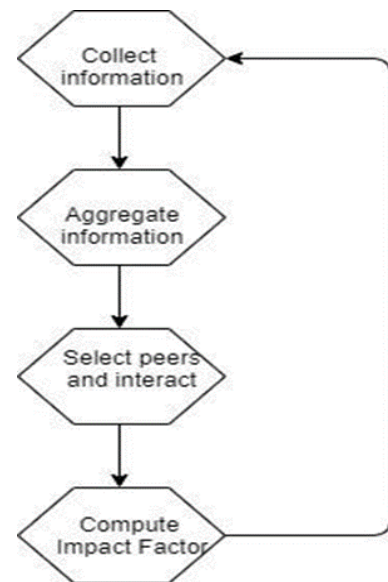


Figure 3. Flow chart Diagram for Endorse Contract

The DApp facilitates the interaction between the users (on their browsers) and the Ethereum Blockchain on the Reputation system [19]. Any node joining the Reputation platform submits his/her application via a DApp which retrieves the public keys from the key store. The keys are used to sign off the data which is then transmitted securely. The DApp then runs the specified smart contract corresponding to the data being transmitted. If the execution is successful, validators pick the transactions and broadcast them to the entire network.

### III. FINDINGS AND DISCUSSION

This section presents the main findings regarding the performance of the reputation system. Specifically, we discuss the system's ability to provide security and privacy, its performance, and the fulfilment of its requirements.

#### A. Security and privacy

Reputation systems can form basis for solutions in the broken online interaction marketplaces. However, as other studies have shown, these systems are vulnerable to different attacks including Sybil attacks, whitewashing attacks, free-rider attacks, and Denial of service attacks [8]. We analyzed our system based on how endorsers' messages get stored on the Blockchain, and how the reputation system computes the impact factor. Whereas our system does not address this issue, available data regarding an entity can be used to determine malicious nodes in the Blockchain ecosystem.

Suppose the reputation system has four nodes ( $y = \{a, b, c, d\}$ ). Here, the reputation platform maintains two sets of information for each entity (one consisting of a list of endorsers, while the other consisting a list of endorsees). If  $m$  is the list of endorsers and  $n$  is the list of endorsees for an entity  $y$ , then, the intersection of  $m$  and  $n$  offers clues on the common entities in the sets. If the intersection set is the same for both the endorsers' list and endorsees' list, then it is a likely indicator that these actors are malicious and wants to interfere with the computation of impact factor.

Determining the actors that could be colluding in a system of certain size assumes is NP-complete problem that can only be solved using heuristics approaches. Such a problem is likely to become complex when more nodes join the network [31]. Essentially, the more the actors join the reputation system, the more the size of the network expands and so does the computational complexity. Since the platform is implemented using Ethereum, the high cost of gas (gas is mentioned just here. Is it correct?) for each transaction does not incentivize participants to act maliciously on the network [15].

#### B. Fulfillment of requirements

A front-end web app prototype was developed to help test the various contract functions on the Blockchain. Manual testing was conducted using truffle IDE and ganache on the local network. Smart contract should be implemented in the form of conditions (checking to ensure all the necessary pre-conditions such as the caller of the function), actions (in form of events and functions) and interactions if such a code is eliminating the reentrancy errors [32],[33]. Our system used the same approach when writing the smart contracts as shown by a snippet function below:

##### Algorithm 1 Verification Function

---

```

1. function joinNetwork()
2.   if the user registered on the network
3.   {
4.     Record the sender's name and id
5.     add new entity to the existing members update the
       member's list
6.   }
7.   else
8.   {Do not join the network} Return

```

---

As shown in the code, the *joinNetwork ()* function first checks to find out whether a node has been registered on the network or not. If the user has not registered, then the function halts. The function progresses with the computation process only if the user has been registered. In this case, the function proceeds to deal with the actions (changing the status of the registered nodes and incrementing the registered nodes).

To ascertain the fulfilment of requirements, simulation was performed with the help of an interaction graph that offered clues regarding the performance of the network. The code below illustrates the various graphical interactions between nodes:

##### Algorithm 2 Graphical Interaction between nodes

---

```

1. function getProfile()
2. {
3.   For each node:
4.     Compute outDegree
5.     Compute used_Power
6.     Compute outConns
7.     Compute inDegree
8.     Compute receivedPoints Return:
9.     outDegree ,used_Power ,outConns ,inDegree
10.    ,receivedPoints ,inConns)
11. }

```

---

The reputation model was then applied to nodes in the interaction graph nodes, and their impact factor calculated based on incoming and outgoing connections ( $E_g$  and  $E_r$ ). Each node on the Ethereum rated one other on a scale of -2 (representing total distrust) to +2 (representing total trust). To provide more relevant findings, the interaction graph was modelled to incorporate only edges that had a rating of +2 with no negative edges in the simulation model. Out of 5000 nodes, 240 edges were marked as positive edges. The information available for each node included source, rating, target, and timestamp which formed the basis for the endorsement system. The direction of endorsement system was based on the source and target datasets while the timestamp data gave hints regarding the order of transactions. The graph below summarizes the distributions of both incoming and outgoing graph connections:

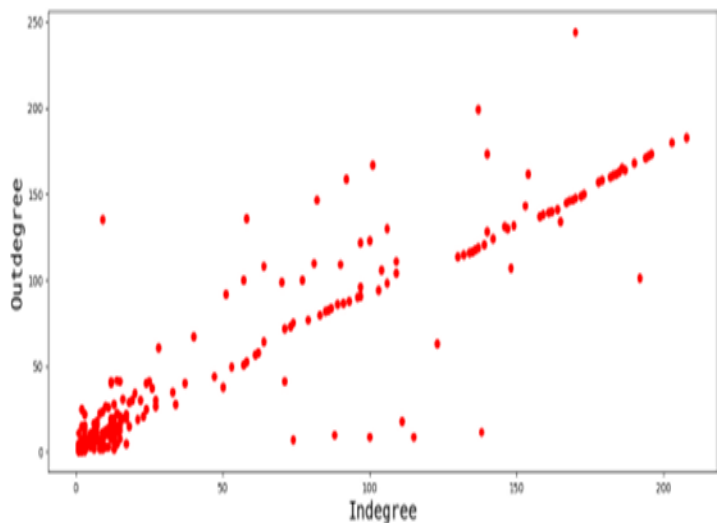


Figure 4: Distribution of incoming and outgoing connections

The simulation model provided basis for collecting datasets about the endorsement system and analyzing it in a manner like a net flow diagram [17]. This information was useful in detecting anomalies within the endorsement system. As evident in figure 1, there were no anomalies with the model which meant the model could be used to compute the impact factor for each node.

The Impact Factor (IF) parameter computed based on  $E_g$  and  $E_r$ . Out of 5000 nodes, 3800 nodes (76%) had an IF of 0. On further examination of these nodes, we found only one incoming or outgoing links (both  $E_g$  and  $E_r$  had values of 1). This makes sense according to our model because a node can only make an impact on the ecosystem if it has more than one connection. In our case, an IF of 0 is expected because it represents the initial node and does not in any way suggest that it is untrustworthy in the system. Based on these findings, it is appropriate to conclude that 76% of the nodes in our network were new users. The remaining nodes (12) had considerable IF scores.

There were no nodes with IF more than 1 that had an accumulated RPTs of 0. If this were the case, then their IF would still be 0 because RPTs is a significant contributor to the reputation platform. Our model shows that it is possible for some nodes to have a maximum possible ratio (which in this case is 1) but still with a low IF. This is true because, as we had mentioned earlier,  $E_g$  and  $E_r$  alone do not contribute to the overall IF on the network. The table below provides some results for selected nodes that had low IF:

Label	$E_g$	$E_r$	Ratio	RPTs	IF
3	2	2	1	0	0
15	5	3	6	1	0.2
41	4	3	0.75	2	0.3

Table 1: Selected nodes with low IF

The results in the above table clearly demonstrate that the ratio between  $E_g$  and  $E_r$  is not the key contributor to the IF in the network. Consequently, it is hard to speculate if a particular entity in the network would have a high IF by merely using the ratio.

#### IV. CONCLUSION

Blockchain adds an extra layer of trust and security to online interactions and can provide a much-needed reputation in online interaction platforms. In this paper, we have presented a trusted, decentralized reputation model based on Ethereum Blockchain. Our platform can help enforce reputation and trust in online interactions. The IF score computed for an entity in the reputation model is a mark of its trustworthiness in the ecosystem. Using this score, any user intending to participate in online interaction with the other party can verify the true identities and proof that the parties are who they claim to be.

While our model is one of the first attempts at leveraging Blockchain to infer trustworthiness of nodes based on peer-to-peer interactions and the computation of IF, it does not in any way challenge existing cryptographic-based reputation systems. Limited time and resources did not allow us to

explore the existing cryptographic-based reputation systems that take into consideration all the factors while computing the impact factor. Further research should be conducted to provide an in-depth overview of current issues in reputation including but not limited to use of Eigen trust systems in reputation systems, application of Anomaly Detection Algorithms to thwart malicious behavior in reputation systems, extend the capabilities of reputation platforms to deal with user accounts and emails.

#### REFERENCES

- [1] 'Digital 2019: Global Digital Overview', *DataReportal – Global Digital Insights*. [Online]. Available: <https://datareportal.com/reports/digital-2019-global-digital-overview>. [Accessed: 06-Oct-2019].
- [2] A. Gonzales, 'The contemporary US digital divide: from initial access to technology maintenance', *Inf. Commun. Soc.*, vol. 19, no. 2, pp. 234-248, 2016.
- [3] Setti, Sunil, and Anjar Wanto. "Analysis of Backpropagation Algorithm in Predicting the Most Number of Internet Users in the World." *Jurnal Online Informatika* 3.2 (2019): 110-115.
- [4] J. K. Rout, S. Singh, S. K. Jena, and S. Bakshi, 'Deceptive review detection using labeled and unlabeled data', *Multimed. Tools Appl.*, vol. 76, no. 3, 3819.
- [5] W. Liu, J. He, S. Han, and N. Zhu, 'A Method for the Detection of Fake Reviews based on Temporal Features of Reviews and Comments', in *3rd International Conference on Mechatronics Engineering and Information Technology (ICMEIT 2019)*.
- [6] Nakamoto S. Bitcoin: A peer-to-peer electronic cash system. *Decentralized Business Review*. 2008 Oct 31:21260.
- [7] A. Bogner, M. Chanson, and A. Meeuw, 'A decentralised sharing app running a smart contract on the Ethereum blockchain', in *Proceedings of the 6th International Conference on the Internet of Things*, 2016, pp. 177–178.
- [8] Shanaev S, Shuraeva A, Vasenin M, Kuznetsov M. Cryptocurrency value and 51% attacks: evidence from event studies. *The Journal of Alternative Investments*. 2019 Dec 31;22(3):65-77.
- [9] Hernan SV, inventor; Microsoft Technology Licensing LLC, assignee. Authentication using proof of work and possession. United States patent applicationUS14/486,864.2016March17.

- [10] Bentov I, Gabizon A, Mizrahi A. Cryptocurrencies without proof of work. In International conference on financial cryptography and data security 2016 Feb 22 (pp. 142-157). Springer, Berlin, Heidelberg
- [11] Nguyen CT, Hoang DT, Nguyen DN, Niyato D, Nguyen HT, Dutkiewicz E. Proof-of-stake consensus mechanisms for future blockchain networks: fundamentals, applications, and opportunities. IEEE Access. 2019 Jun 26;7:85727-45.
- [12] Tosh, Deepak, et al. "CloudPoS: A proof-of-stake consensus design for blockchain integrated cloud." *2018 IEEE 11th International Conference on Cloud Computing (CLOUD)*. IEEE, 2018.
- [13] Watanabe, H., Fujimura, S., Nakadaira, A., Miyazaki, Y., Akutsu, A., & Kishigami, J. (2016, January). Blockchain contract: Securing a blockchain applied to smart contracts. In *2016 IEEE international conference on consumer electronics (ICCE)* (pp. 467-468). IEEE.
- [14] Nugent, Timothy, David Upton, and Mihai Cimpoesu. "Improving data transparency in clinical trials using blockchain smart contracts." *F1000Research* 5 (2016).
- [15] Cong, Lin William, and Zhiguo He. "Blockchain disruption and smart contracts." *The Review of Financial Studies* 32.5 (2019): 1754-1797.
- [16] Dennis R, Owen G. Rep on the block: A next generation reputation system based on the blockchain. In 2015 10th International Conference for Internet Technology and Secured Transactions (ICITST) 2015 Dec 14 (pp. 131-138). IEEE.
- [17] Fichtner, Wolfgang, Qiuting Huang, Bernd Witzigmann, Hubert Kaeslin, Norbert Felber, and Dölf Aemmer. "IIS Research Review 2004." *IIS Research Review* (2004).
- [18] Y. Zhang, C. Xu, J. Ni, H. Li, and X. S. Shen, 'Blockchain-assisted public-key encryption with keyword search against keyword guessing attacks for cloud storage', *IEEE Trans. Cloud Comput.*, 2019.
- [19] G. Liang, S. R. Weller, F. Luo, J. Zhao, and Z. Y. Dong, 'Distributed blockchain-based data protection framework for modern power systems against cyber attacks', *IEEE Trans. Smart Grid*, vol. 10, no. 3, pp. 3162-3173,
- [20] S. Eskandari, S. Moosavi, and J. Clark, 'SoK: Transparent Dishonesty: front-running attacks on Blockchain 2019.
- [21] Cachin C. Byzantine faults. In Concurrency: The Works of Leslie Lamport 2019 Oct 4 (pp. 67-81).
- [22] Gramoli, Vincent. "From blockchain consensus back to Byzantine consensus." *Future Generation Computer Systems* 107 (2020): 760-769.
- [23] L. Luu, V. Narayanan, K. Baweja, C. Zheng, S. Gilbert, and P. Saxena, 'Scp: A computationally-scalable byzantine consensus protocol for blockchains', *SCP*, vol. 20, no. 20, p. 2016, 2015.
- [24] S. Tamang, *Decentralized Reputation Model and Trust Framework Blockchain and Smart contracts*. 2018.
- [25] Fernández Anta, A., Rimba, P., Abeliuk, A., Cebrian, M., Stavrakakis, I., Tran, A. B., & Ojo, O. (2019). Miner Dynamics on the Ethereum Blockchain.
- [26] King, Sunny, and Scott Nadal. "Ppcoin: Peer-to-peer crypto-currency with proof-of-stake." *self-published paper, August 19.1* (2012).
- [27] W. Wang, 'A Vision for Trust, Security and Privacy of Blockchain', in *International Conference on Smart Blockchain*, 2018, pp.93-98.
- [28] M. Vukolić, 'The quest for scalable blockchain fabric: Proof-of-work vs. BFT replication', in *international workshop on open problems in network security*, 2015, pp.112-125.
- [29] C. Kaligotla and C. M. Macal, 'A generalized agent based framework for modeling a blockchain system', in *Proceedings of the 2018 Winter Simulation Conference*, 2018, pp.1001-1012.
- [30] Berger TP, Gueye CT, Klanti JB. A NP-complete problem in coding theory with application to code based cryptography. In International Conference on Codes, Cryptology, and Information Security 2017 Apr 10 (pp. 230-237). Springer, Cham.
- [31] M. Prates, P. H. Avelar, H. Lemos, L. C. Lamb, and M. Y. Vardi, 'Learning to Solve NP-Complete Problems: A Graph Neural Network for Decision TSP', in *Proceedings of the AAAI Conference on Artificial Intelligence*, 2019, vol. 33, pp. 4731-4738.
- [32] Y. Hou, X. Zhao, Q. Li, J. Chen, Y. Li, and Z. Zheng, 'Solving large-scale NP-Complete problem with an optical solver driven by a dual-comb "clock"', in *CLEO: Science and Innovations*, 2019, pp. SF1M-2.
- [33] Sultan K, Ruhi U, Lakhani R. Conceptualizing blockchains: characteristics & applications. arXiv preprint arXiv:1806.03693. 2018 Jun 10.
- [34] DelGaudio CI, Hicks SD, Houston WM, Kurtz RS, Hanrahan VA, Martin Jr JA, Mummey DP, Murray DG, Prince JE, Pritsky RR, Rauch DC, inventors. Method and system for network connectivity migration management. United States patent US 9,928,480. 2018 Mar 27.
- [35] Efanov, Dmitry, and Pavel Roschin. "The all-pervasiveness of the blockchain technology." *Procedia computer science* 123 (2018): 116-121.

# An Observation on Residential Complexes as a New Housing Typology in Post-Socialist Tirana

Edmond Manahasa  
Department of Architecture  
Epoka University  
Tirana, Albania  
emanahasa@epoka.edu.al

Rudina Kazazi  
Department of Architecture  
Epoka University  
Tirana, Albania  
rkazazi09@epoka.edu.al

**Abstract**— The shift from socialist to post-socialist political system caused an urban expansion of Tirana, which brought the growth of its population. This post-socialist urban context associated with a high demand for dwelling need, the limited construction sites in inner city and the increasing land prizes, pushed the developers to construct residential complexes as a new form of housing in Tirana. This research aims to reveal the features of residential complexes as a new housing typology, that emerged during the post socialist period in Tirana by analyzing typological housing features and the common outdoor spaces. To investigate properly there are selected three cases studies: “Halili”, “Homeplan” and “Kika” residential complexes. The methodology used in the research includes visual documentation, archival research for provision of the drawings and an analysis of typological features and outdoor spaces of the residential complexes. The study revealed that residential complexes in Tirana are featured by dense high-rise apartment blocks or towers, which possess many apartments. The spatial features are characterized low number cores, which in majority are not lit and provide access to large number of apartments. Apart from the housing functions, commercial functions are found in the ground floors, whereas in some cases the business functions are located also in upper floors. The common outdoor spaces are featured by limited green spaces and considerably are usurped by the commercial units of the ground floors.

**Keywords**—residential complexes, housing typology, post-socialist period, Tirana

## I. Introduction on Multifamily Housing in Post-socialist Tirana

The multifamily housing development is one of the major dwelling forms in the post-socialist period Tirana. The post-socialist period housing developments in Tirana are featured by informal housing settlements at the city periphery and high-rise dwellings in the inner city [1]. Although after 90s, due to technological limitations the first multifamily housing were middle-rise apartment blocks, by the years 2000 in a dominant majority they evolved in the form of high-rise apartment blocks or residential complexes.

While for an apartment blocks the needed site was smaller, in the case of residential complexes there was a need for larger sites. Based on that on the case of residential complexes it was easier to be developed in the peri urban zones of the city (out of middle ring), but those zones were lacking the needed infrastructure making the unattractive for the clients. Thus, it was more attractive to develop residential complexes in the inner-city zones which controversially lacked such large sites for development. Taking into consideration the complex issue of land ownership in Albania and the weak urban planning administration the target of developers for residential complexes sites become the public green and sportive spaces (case of “Fusha e zeze” field or “Partizani” club training grounds) or socialist period enterprises which had gone

bankrupt or stopped production [2] (case of ex-mechanical factory Enver Hoxha).

The number of the residential complexes that were developed in peri-urban areas after years 2010, in consistence with the expansion of the city have increased in quantity. The characteristic feature of residential complexes are the high-rise apartment blocks (in many cases tower shaped) aiming maximize the dwelling area. They are arranged at the perimeter of the rectangular sites and in appearance the spaces in between are dedicated for outdoor social activities. However, in most of the cases the ground floors are designed for coffee shops, which easily usurp the public spaces [3].

## II. Case Study on Three Residential Complexes in Post-Socialist Tirana

This research focuses on the analysis of three case studies of residential complexes in Tirana. The analysis includes description of site plans, housing typological features and the common outdoor spaces of residential complexes. The analysis is supported with visual material for each of them, including photos, to plan drawing and technical materials. Firstly, an introduction about the site plan and urban context of the residential complexes, including the position related to the city center is depicted. The housing typological features include architectural composition, spatial features, function distribution and exterior solutions. The common outdoor spaces explain the shared spaces within the sites of the residential complexes, their organization and hardscape and green spaces ratios. Furthermore, it includes observation on the usage these spaces, and the relation of between the dwellers and commercial units located on ground level and, parking accesses or passages.

### A. “Halili” Residential Complex

“Halili” complex is conceptualized as a residential development and business center. It is located in the inner city of Tirana, on the Dibra street within an old neighborhood approximately 820 meters from the city center. Designed by Vladimir Bregu, it was completed in 2004. Although initially it was planned to cover a larger zone (approximately 13,900 sq m<sup>2</sup>), its site is approximately 7200 sq m<sup>2</sup> [4]. The site plot has a triangular-like shape and excluding the “Partizani” high school, which border the complex in its rear part. The design approach was to have a built perimeter on the main streets while enclosing the inner space, giving it private attributes.

Based on that, the perimeter of the land plot is contoured by high rise apartment blocks, that vary in height from a minimum of 6 floor height up to 11 floor height residential towers. The common outdoor spaces are located on the back of the build perimeter, providing the necessary intimacy and



Figure 1. Location of "Halili" residential complex (left), image (middle) and its common outdoor (right).

security from the 3 vehicular streets that surround the complex. The access from the main road to the inner space of the complex is done through the passageways. The ground floor of the complex is an alternation of commercial units, entrances on the upper floor business units and residents.

### A.1. Housing Typological Features

The architectural composition of this residential complex evolves as an interplay of cylindrical, oval, and edgy volumes, giving a dynamic impact. The streetside façade balconies are extruded volumes from the main volume, as consoles their shape varies from floor to floor, creating the illusion of movement in the façade. The corners of the plot are treated with a circular element creating a smooth transition at the junctions, for circulation. The materials used in the grey ceramic tiles, whereas the rounded volumes are clad with curtain wall façade, which due to the large cantilever and complicated geometry produce a high-tech and dramatic effect.

The apartment blocks are designed into two different types and vary in height differing from 6 floor height up to 13 floors. One of them apart from residential functions, also accommodates business ones and ends in a semicircular form and is placed perpendicular to the Dibra street, whereas the other which is placed parallel to the street is used for residential functions and possess circular balconies.

The spatial organization of the apartment blocks is featured by staircase cores that are located on the central parts of the building and do not receive neither daylight nor natural ventilation. The distance of the apartment entrances and cores is short, diminishing also the shared space as hallways. Hallways lack natural light thus are supplied all day and night through the help of artificial lighting. Their width varies from a minimum of 145 cm up to 160 cm. Each staircase feeds a maximum of 4-5 apartments per floor accompanied by two elevators with a carrying capacity of 4 people each.

### A.2. Common Outdoor Spaces

The common outdoor spaces in this residential complex are minimal. Since the complex has underground parking for the residents, there are two ramp entrances, which compromise and reduce the public space. While going through the public space, several parts are occupied by cars parked also on ground level being an obstacle in the circulation of people and in using all of it.

The materials used for the flooring of the public space are a combination of stamped concrete and the red brick applied as fishbone pattern in areas, where urban furniture and trees are located. Greenery presence in the area consists of small to medium deciduous trees, positioned in concrete vases and spread throughout the outdoor space. They do not provide enough shadow during summertime, thus resulting in a hot environment during peak summer hours.

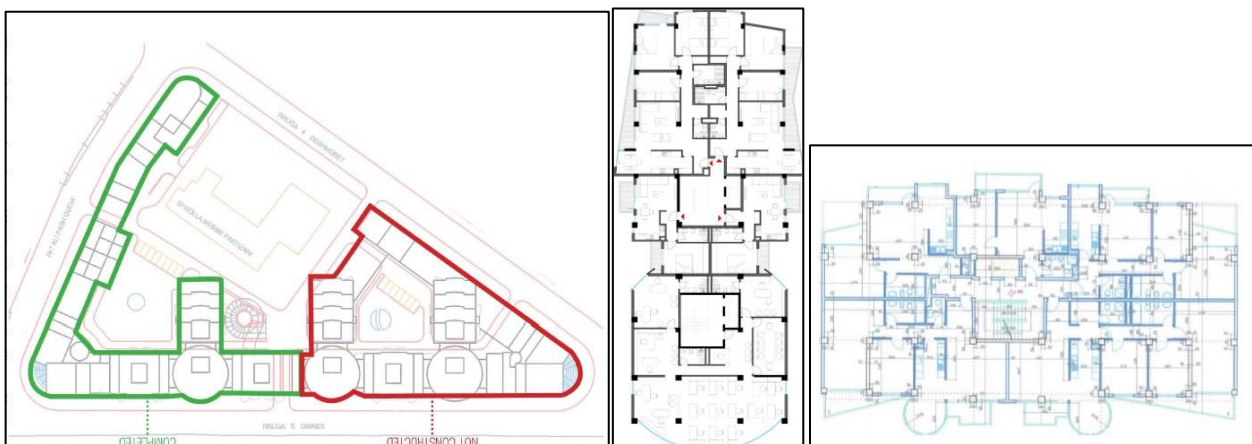


Figure 2. Site plan of "Halili" complex with realized parts in green (left), perpendicular (middle) and parallel housing block (right) normal floor plan schemes

The commercial units located on the ground floor, that look to the inner outdoor common space, mostly are not occupied and the small number of which do have a tenant consists of small mini markets or dry-cleaning activities.

Related to the usage of public space, it is generally a quiet place, although it is not vehicle free. The people using the space are generally old people, or young parents that carry their infants in the strollers spending a bit of quiet time outdoors. The space is populated during the late afternoon hours when the sun is setting, by children playing unattended by their parent, while most of the residents tend to sit at the bars located nearby.

## B. "Homeplan" Residential Complex

"Homeplan" is designed as a residential complex composed of residential and commercial units. It is located in the western part of Tirana, outside the middle ring. It is designed by architect Irina Branko and completed a local company called Kontakt. The "Homeplan" residential complex site plan scheme organization, consists of building volumes placed at the perimeter of a quasi-rectangular site and an inner courtyard in the center. It is approximately 2.4 km from the city center and is positioned on a 6751 m<sup>2</sup> plot. As a construction period it has started in 2010 and ended in 2012. In verticality it reaches up to a maximum of 8 floors high with one underground parking floor. Apart from housing, other facilities present within "Homeplan" complex consist of service units as bars, restaurants, or supermarkets, which are placed on the ground floor [5].

number of apartments accessed from the same core varies from five to seven. The cores and the corridors are not lit in majority of the cases. Related to the commercial units and their location, the ground floor as the easiest space to directly having an access from the main road are in their function. The entrances of the residents are positioned in such a way to not interfere and occupy a valuable front-line commercial unit, but they are positioned in the inner part of the passageways. All the entrances are numbered, to ease the directionality and orientation of the dwellers.

The exterior of this residential complex due to the innovative materials and preciseness in realization technique evolve, thus the difference can be clearly distinguished among vicinity buildings. The façades of this residential complex are coated by raw stucco finishing. The finishing colors are sensitive and are pinkish pastel color, light grey and white.

### B.2. Common Outdoor Spaces

The common outdoor spaces of the "Homeplan" residential complex is a central square plaza arranged in the inner courtyard. Although this space is semi closed and introverted, it is usable and accessible also for the dwellers off complex. The public space has a coverage area of approximately 1200 square meters, where urban furniture, greenery and water features are integrated.

The materials used in the inner courtyard consists of grey stamped concrete with a stone like pattern. Minimal green areas are designed in rectangular upraised volumes 50 cm

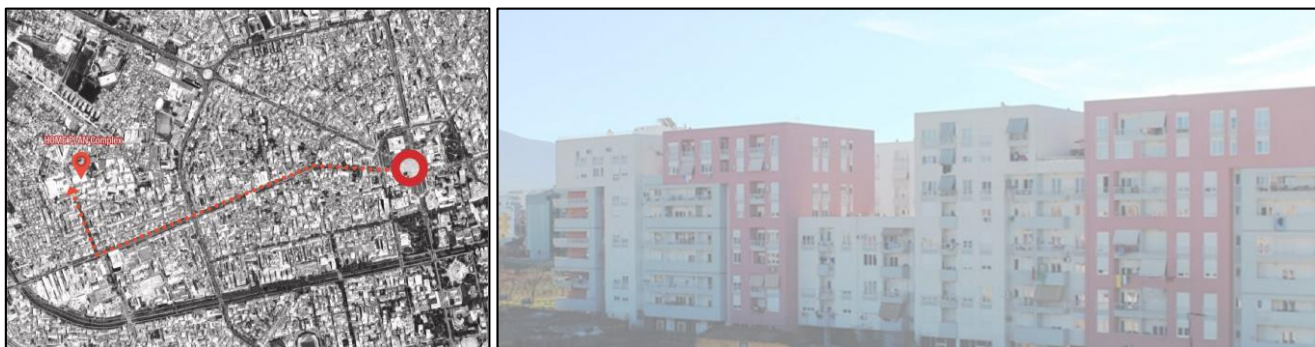


Figure 3. "Homeplan" residential complex location (left) and its image (right)

### B.1. Housing Typological Features

There are seven apartment blocks in this residential complex. The built volumes rise as extruded volumes, while some of them do not continue to the last floor but reach to a maximum of 5 floor high. In this way the roof is converted to an open veranda for the users of the 6th floor adding value to the above apartments as well as liberating the massive impact of one whole volume. Under the lower volumes in floor height, there are passageways on ground level of the building. The width of the buildings varies from 14.8 meters to 20 meters. On the outer perimeter the balconies and loggias are cantilevered, while on the inner perimeter they stay within the built perimeter of the building.

The spatial organization of apartment blocks is planned around the cores which are placed at the center of respective block. The cores consist of a staircase and two elevators. The

from the ground level, which offset from the inner building perimeter. The inner part of these volume is designed as a hilly - like terrain, planted in grass, bushes and high trees. Urban furniture is integrated within the perimeter of the volumes. The urban furniture partly is designed with wooden elements, which create a warmer and more inviting element than concrete.

The services located on the ground floor are bars and cafeterias, which usurp of the common outdoor space for their private purposes. This results in a limited area for the user, who does not want to use the cafeterias nearby. Based on the observations we conducted on site, the large number of coffee

shops push the young parents to sit in there, to keep their children under their supervision.

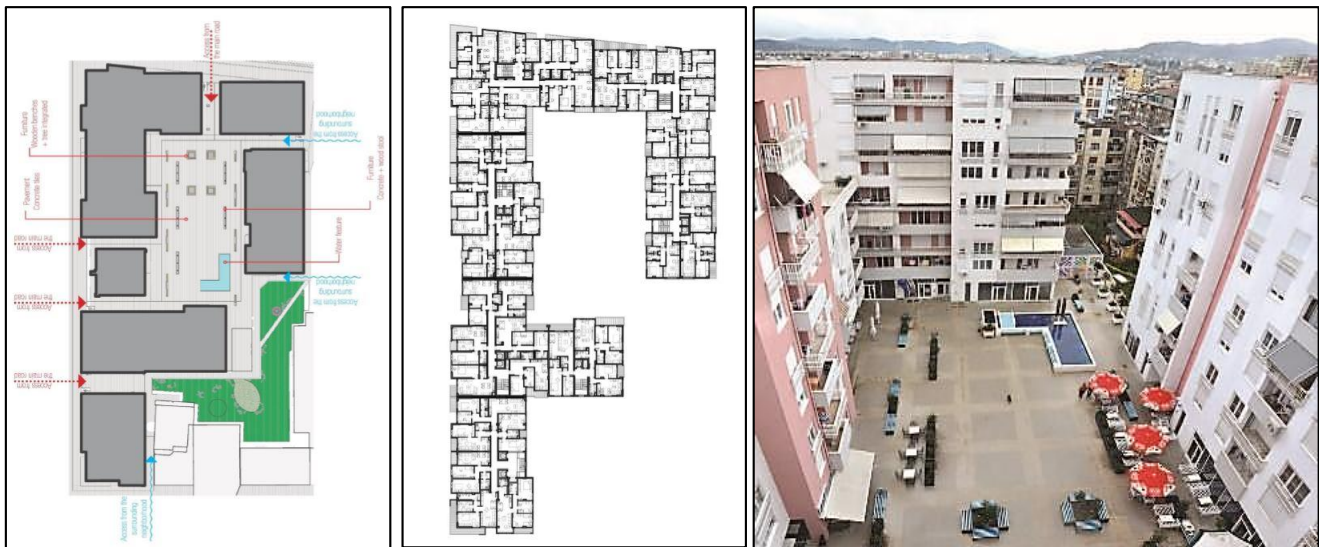


Figure 4. "Homeplan" site plan (left), normal floor plan scheme (middle) and image of common outdoor spaces (right)

### C. "Kika" Residential Complex

The residential complex "Kika" is as a mixed-use housing development. It is located in the south-western part of Tirana, outside of middle ring approximately 1.8 km from the city center in. Priorly during the socialist period, the site was used as a greenhouse. Designed by Italian Atena studio it is developed in a site of 40,000 m<sup>2</sup> site and is planned in three large divisions. The three divisions belong to a previously larger area which was developed through the known "Le Serre masterplan design" [6].

The site plan of residential complex is featured by housing blocks which are planned at the perimeter of each of the divisions, whereas the common outdoor spaces are designed in the form of central courtyards.

The volumetric composition of these large mass housing development is featured by dense tower buildings, which are fragmented vertically have large cantilevers and, in some cases, connects to each other with bridges. Beyond housing functions other facilities present within the residential complex are restaurants, commercial areas, shopping centers, spa, cinemas, and auditoriums. The construction process started in 2005 and is supposed to finish in 2020 [7].

### C.1. Housing Typological Features

To see in a closer way, the housing features we will examine the 3<sup>rd</sup> division of the residential complex. This division has a site in a trapezoidal shape, which has influenced the housing design to have a sharp angular façade.

The buildings are mainly oriented on a north south axis, meanwhile the apartments are oriented towards east and west directions. This orientation is favorable in maintaining a cool inner environment in summer and a lower need for heating in winter resulting in a considerable amount of energy saved. Since the blocks are linked between each other and their proximity at several points may cause shading or privacy problems, the solution was to fragment the height into several, starting from a minimum of 6 floor height buildings, 7,8 and reaching a maximum of 9 floor height in the 3<sup>rd</sup> division. The width of the building blocks varies from 23 meters up to 25 meters. In this residential complex is observed a usage of extreme cantilever volumes. Beyond those cantilevered parts which can reach up to 7-8 meters, within the complex there are a set of linking bridges. The large, cantilevered volumes are designed in steel structure in order to provide the statical stability.

The spatial organization of the apartment blocks, include minimum number cores and linear long corridors. The



Figure 5. Image of "Kika" residential complex (left) and common outdoor spaces (right)



number of apartments accessed from one core varies from five to eleven. In some cases, certain apartment block is not designed with a core, and they are connected via bridges to the other block core. Apart from the ground floor is allocated for commercial activities, in the angular part in difference from the previous block the first floor is totally dedicated to commercial activities, shifting the apartment units to the second floor up to the ninth.

The façade of this residential complex is covered in grey plaster, giving the blocks a neutral look. Meanwhile the outer and inner façade is fragmented, resulting in a much lighter volume and game of solids and voids. The usage and interplay of vertical windows on the perimeter of the block serve to break the horizontalness of the built environment. Meanwhile, the inner façade differs from the outer, is a more porous one, offering more transparency and light to the inner environment. The inner façade consists of continuous loggias that are divided by walls between 2 apartments, providing the necessary privacy between the neighbors.

### C.2. Common Outdoor Spaces

The common outdoor spaces of the residential complex are a result of the arrangement of the building blocks in the perimeter of the plots, resulting in an introverted courtyard. The linkage between these inner spaces and the surrounding neighborhood is done through several punctuations in the built volumes, in forms of passages that also serve as controlled entrances to these spaces. Each of these public spaces have two or three of these punctuations that either connect them with the neighborhood or the sites with each other. The built volumes reach up to a maximum of 9 floor height, despite that the distance among them provides the opportunity for the inner environment to receive plenty of natural daylight through the day. The materials used in the inner spaces consists of grey stamped concrete stone like pattern [6].

The green spaces are abundant and consist of two orthogonal islands in the inner court and another smaller one in the eastern edge of the site. They are designed as an upraised volume from the ground level at approximately 50 cm. The topography of the green spaces is designed as a hilly-like terrain and is planted in grass, bushes, and high trees. The perimeter of the green spaces is designed as urban sitting furniture consisting of concrete bordure and wooden elements, providing a warmer environment.

### III. Concluding Remarks

To conclude it can be said that the residential complexes are a new form of dense mass housing that have emerged in considerable quantity during the post-socialist period in Tirana. The increasing number of this housing form is related to the capitalist economic reality, based on which the developers aim to maximize the housing construction profit and the purchasing power of the citizens is not high. Due to this context, the residential complexes that were subject of this study in majority are featured by dense and high-rise tower like apartment blocks.

Regarding their spatial organization their cores consist of stairs which are unlit and provide access to large quantity (five to twelve) of apartments. In some cases, there are observed also long linear corridors. The exterior of this housing typology reflects the quality of contemporary construction technology and in many cases, has played an avant-gardist role using curvilinear shapes and curtain wall façades.

The common outdoor spaces generally are designed in the form of inner courts. Due to the commercial activities in the ground floors of the housing blocks and the lack respective legal framework needed to regulate the usage of the common outdoor spaces, predominantly they are subject of usurpation by the owners of these activities. The green spaces are present in the inner courtyards, although they are in minority compared to overall space.



Figure 6. Kika residential complex in green within Le Serre master plan (left), "Kika" 3rd division site plan (middle) and normal floor plan (right)

## References

- [1] Manahasa Edmond, Manahasa Odeta, “Defining urban identity in a post-socialist turbulent context: The role of housing typologies and urban layers in Tirana”, *Habitat International*, Volume 102, 102202, 2020.
- [2] Manahasa, Edmond, Ozsoy, Ahsen, “Place attachment to a larger through a smaller scale: Attachment to city through housing typologies in Tirana”. *Journal of Housing and the Built Environment*, 35(1), 265–286, 2020.
- [3] Manahasa, Edmond, “Place attachment as a tool in examining place identity: A multilayered evaluation through housing in Tirana”, PhD dissertation, Istanbul Technical University, 2017.
- [4] Halili construction archive, July 2018.
- [5] Kontakt archive, July 2018.
- [6] Kika Construction archive, July 2018
- [7] <http://www.atenastudio.it/sitoweb/project.php>

# Active Control Of In-Wheel Motor Electric Vehicle Suspension Using The Half Car Model

Abbas Hussein Hammood  
Department of Mechatronics Engineering  
Istanbul Okan University  
Istanbul, Turkey  
abbhammood@stu.okan.edu.tr

Başar Özkan  
Department of Mechanical Engineering  
Istanbul Okan University  
Istanbul, Turkey  
basar.ozkan@okan.edu.tr

**Abstract**— In recent years, electric vehicles are becoming more mainstream. Even though in-wheel electric motor vehicles have been developed as prototype vehicles, they have not become common yet. One of the reasons for this is that when an electric motor is placed in the wheel it becomes heavier, which tends to worsen the road holding properties of the vehicle. Active suspensions are currently used in some high-end vehicles to improve passenger comfort, vehicle handling and road holding. In this paper an active suspension is used show that it is possible to mitigate the road holding problems caused by in-wheel electric vehicles. The ground-hook method is used for the control strategy. First, a quarter car model is used to show that adding a weight to the wheel does actually decrease the road holding performance of the vehicle. Afterwards, the active suspension is shown to be able to improve the road holding properties of the vehicle to acceptable levels. The simulations are also repeated with a half car model. They also show that increasing the tire mass worsens the road holding performance of the vehicle. However, this bad performance can be reversed by the ground-hook active suspension. Simulations also show that the ground-hook controller worsens the passenger comfort level in the vehicle.

**Keywords**—active suspension, road holding, ground-hook, electric vehicle, in-wheel motor

## I. Introduction

The suspension system is the one of the most important part of an automobile that isolate the vehicle from road shocks, vibrations and provide comfort effect to the occupant [1][2]. Automotive suspensions are divided into three forms namely passive, semi-active, and active suspension system. Passive suspensions are always used and continuous improvements have been made by research. It is impossible to get both ride comfort and road holding demands in the same time for passive suspension car. Passive suspension systems are the most favorite and are widely used because of their low cost and high reliability. This type of system is considered as open loop system [2]. Passive suspension consists of conventional spring (the spring is pressed and stretched to absorb the wheel movement) and damper which is a shock absorber that works on the vibration motion of the vehicle. The main aim of using damper is to slow down and minimizing the vibration magnitude caused by the road. The damper connected in parallel with spring which was fixed and it is impossible to change externally by any signal [3][4]. So, it should need a spring which can be stiff and soft simultaneously [3][4][5]. Researchers have made a lot of improvements over the years and most of these experts think that the passive suspension are hard to be improved. A ground-hook control is one of the control strategies applied to the automotive suspension. A ground-hook controller is used to improve the road holding for both quarter and half car models. This controller method

supposes that it is virtually connecting a damper between the unsprung mass and the ground [6]. The ground-hook main task is to reduce the vertical displacement of the tire and keep the ground-tire contact force in a narrow area as possible to the main value [7]. Ground-hook has improvement when compared with passive suspension model. The use of the half car model to show the effects of the ground-hook controller is the main contribution of this paper. The goal is to design control strategy namely ground-hook controller to improve road handling for the vehicle.

## II. Quarter Car Model

### A. Quarter Car Model Passive Suspension System

Quarter Car model is the most popular model that is used in analysis of automotive suspensions and design. The main reason to use this model is that it is simple, it can give reasonably accurate information and predict lot of important properties about the full car, Figure 1 shows a passive suspension system for a quarter car in which the wheel is connected to the body of the car by passive parameters (spring and damper), and the tire is represented as a spring and the damper of the tire is neglected. The quarter car structural model involved mass of car ( $m_s$ ) and mass of tire ( $m_u$ ). There are three vertical displacement types included in quarter car, the vertical displacement for the mass of car ( $z_s$ ), the vertical displacement for the mass of tire ( $z_u$ ) and the vertical displacement for the road ( $z_r$ ).

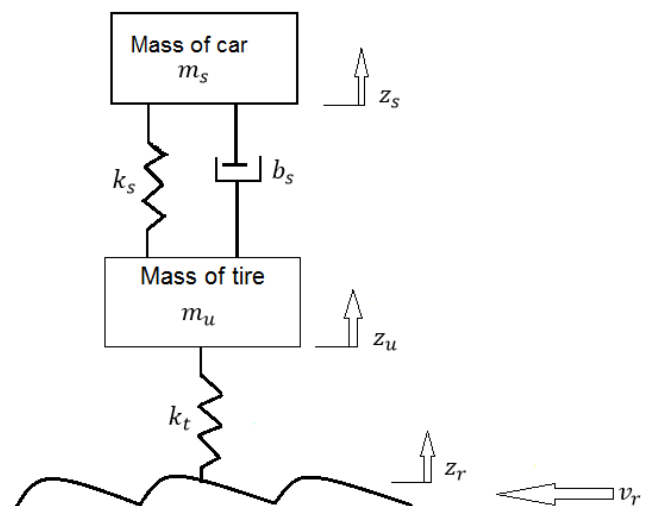


Figure 1. Passive suspension system for a quarter car model

### B. Quarter Car Model Forces

$$\sum f = ma \quad (1)$$

The variables are: ( $f_s$ ) force of spring, ( $f_b$ ) force of damper, ( $f_t$ ) force of tire, ( $m_s$ ) mass of car, ( $m_u$ ) mass of tire, ( $\ddot{z}_s$ ) the acceleration for mass of car, ( $\dot{z}_s$ ) the velocity for mass of car, ( $z_s$ ) position for mass of car, ( $\dot{z}_u$ ) the acceleration for mass of tire, ( $\dot{z}_u$ ) the velocity for mass of tire, ( $z_u$ ) position for mass of tire.

$$-f_s - f_b = m_s \ddot{z}_s \quad (2)$$

$$m_s \ddot{z}_s = -k_s(z_s - z_u) - b_s(\dot{z}_s - \dot{z}_u) \quad (3)$$

$$\ddot{z}_s = -\frac{k_s}{m_s}(x_1) - \frac{b_s}{m_s}(x_2) + \frac{b_s}{m_s}(x_4) \quad (4)$$

$$f_s + f_b - f_t = m_u \ddot{z}_u \quad (5)$$

$$m_u \ddot{z}_u = k_s(z_s - z_u) + b_s(\dot{z}_s - \dot{z}_u) - k_t(z_u - z_r) \quad (6)$$

$$\ddot{z}_u = \frac{k_s}{m_u}(x_1) + \frac{b_s}{m_u}(x_2) - \frac{b_s}{m_u}(x_4) - \frac{k_t}{m_u}(x_3) \quad (7)$$

From eq. (4) and eq. (7) we will get the matrix of A

$$A = \begin{bmatrix} 0 & 1 & 0 & -1 \\ -\frac{k_s}{m_s} & -\frac{b_s}{m_s} & 0 & \frac{b_s}{m_s} \\ 0 & 0 & 0 & 1 \\ \frac{k_s}{m_u} & \frac{b_s}{m_u} & -\frac{k_t}{m_u} & -\frac{b_s}{m_u} \end{bmatrix} \quad (8)$$

$$L = \begin{bmatrix} 0 \\ 0 \\ -1 \\ 0 \end{bmatrix} \quad (9)$$

### C. In-Wheel Motor Electric Vehicle

The In-Wheel Motor Electric Vehicle (IWM EV) is one of the common types of the electric vehicles, the suspension system of (IWM EV) is almost doubled the mass of the wheel by adding the mass of an electric motor to the mass of the wheel. The in-wheel motor (IWM) is placed in the wheel empty space [8]. It was known there is relative relationship between the body vibration and the mass of the wheel, when the mass of the wheel increased the body vibration also increased that's affecting on passenger comfort [8]. An (IWM EV) has extra wheel mass because of the electric motor mass added to the mass of wheel, which lead to discomfort for the passenger and also not safety on the road [9]. The increase of unsprung mass cause the vehicle to get worse riding comfort and handling stability [10,11]. Which affecting on the tire-road contact [11].

$$(m_u + m_i) = m_t \quad (10)$$

$$n_1 = m_t \ddot{z}_u \quad (11)$$

$$n_1 = f_s + f_b - f_t \quad (12)$$

$$n_2 = k_s(z_s - z_u) + b_s(\dot{z}_s - \dot{z}_u) \quad (13)$$

$$n_3 = -k_t(z_u - z_r) \quad (14)$$

$$n_1 = n_2 + n_3 \quad (15)$$

$$\ddot{z}_u = \frac{k_s}{m_t}(x_1) + \frac{b_s}{m_t}(x_2) - \frac{b_s}{m_t}(x_4) - \frac{k_t}{m_t}(x_3) \quad (16)$$

From eq. (4) and eq. (16) we will get the matrix of A

$$A = \begin{bmatrix} 0 & 1 & 0 & -1 \\ -\frac{k_s}{m_s} & -\frac{b_s}{m_s} & 0 & \frac{b_s}{m_s} \\ 0 & 0 & 0 & 1 \\ \frac{k_s}{m_t} & \frac{b_s}{m_t} & -\frac{k_t}{m_t} & -\frac{b_s}{m_t} \end{bmatrix} \quad (17)$$

The L matrix is shown in eq. 9.

### D. Ground-hook Control

A ground-hook control introduces a damper connected virtually to the ground as modelled in Figure 2, ( $c_{grad}$ ) connected between the mass of the wheel and the fixed imaginary frame on the ground. The ( $c_{grad}$ ) is a ground-hook damping coefficient. The damper ( $c_{grad}$ ) is connected to  $m_u$  (the mass of tire) instead of  $m_s$  (the mass of the car). The ground-hook controller improves vehicle road holding by minimizing the upper and lower peaks of wheel displacement and tire deflection [12].

$$f_{grad} = c_{grad}(\dot{z}_u - \dot{z}_r) \quad (18)$$

$$B = \begin{bmatrix} 0 \\ \frac{1}{m_s} \\ 0 \\ -\frac{1}{m_t} \end{bmatrix} \quad (19)$$

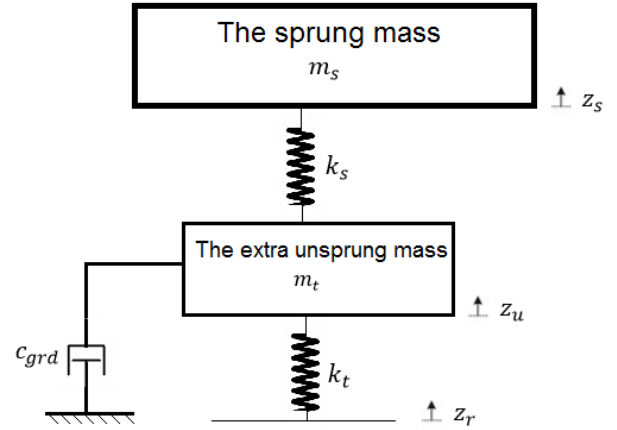


Figure II. In-wheel motor electric quarter vehicle model with groundhook controller

The real state space of vehicle is

$$\dot{x} = Ax + Bu \quad (20)$$

$$y = Cx + Du \quad (21)$$

A is the state matrix, B is the input matrix, C is the output matrix and D is the direct transmission matrix.

$x$  is state vector and  $u$  is control vector

Where:

$$x_1 = z_s - z_u \quad (22)$$

$$x_2 = \dot{z}_s \quad (23)$$

$$x_3 = z_u - z_r \quad (24)$$

$$x_4 = \dot{z}_u \quad (25)$$

$x_1$  is the car suspension deflection,  $x_2$  is the sprung mass velocity,  $x_3$  is the tire deflection and  $x_4$  is the unsprung mass velocity.

The ground-hook controller is compared with passive suspension for both (normal vehicle and in-wheel motor electric vehicle). The simulations were carried out in MATLAB/Simulink.

### Parameter for normal quarter vehicle model passive suspension system

Mass of car  $m_s=250$  kg, mass of tire  $m_u=40$  kg, the spring stiffness  $k_s=16000$ , the tire stiffness  $k_t=160000$ , the damper  $b_s=1000$ .

### Parameter for in-wheel motor quarter electric vehicle model passive suspension system

Mass of car  $m_s=250$  kg, mass of tire  $m_u=40$  kg, mass of in-wheel motor  $m_i=45$  kg, the spring stiffness  $k_s=16000$ , the tire stiffness  $k_t=160000$ , the damper  $b_s=1000$ .

( $m_i$ ) is the mass of in – wheel motor.

### Parameter for in-wheel motor quarter electric vehicle with ground-hook controller

Mass of car  $m_s=250$  kg, mass of tire  $m_u=40$  kg, mass of in-wheel motor  $m_i=45$  kg, the spring stiffness  $k_s=16000$ , the tire stiffness  $k_t=160000$ , ground-hook damping coefficient ( $c_{grad}$ ) =5000.

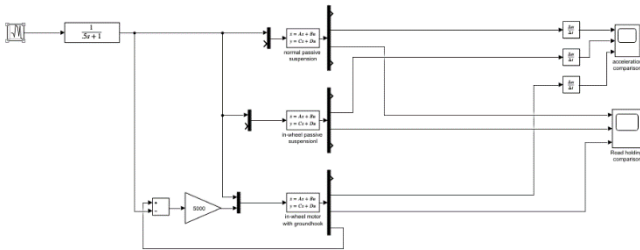


Figure III. MATLAB simulation for quarter vehicle model with normal passive, in-wheel motor passive suspension system and in-wheel motor with ground-hook controller

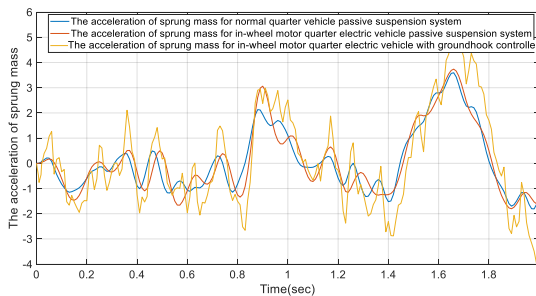


Figure IV. Comparison in acceleration of sprung mass for quarter vehicle between normal passive, IWM EV passive and IWM EV with ground-hook controller

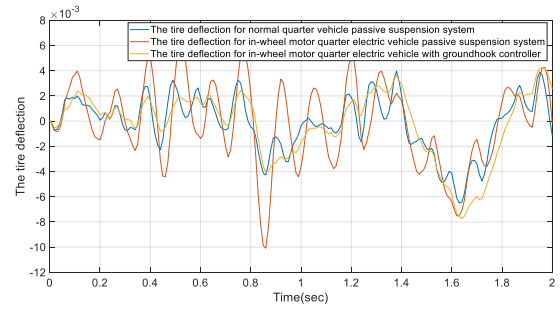


Figure V. Comparison in tire deflection for quarter vehicle between normal passive, IWM EV passive and IWM EV with ground-hook controller

### E. Ground-hook and Sky-hook Control Comparison

A sky-hook control introduces a damper connected virtually to the sky, ( $c_{sky}$ ) connected between the mass of the car and the fixed imaginary frame to the sky. The ( $c_{sky}$ ) is a sky-hook damping coefficient. The damper ( $c_{sky}$ ) is connected to  $m_s$  (the mass of the car). The sky-hook controller improves comfort of the occupants inside the car by minimizing the upper and lower peaks of car body displacement and acceleration of sprung mass[13].

$$f_{sky} = -c_{sky}(\dot{z}_s) \quad (26)$$

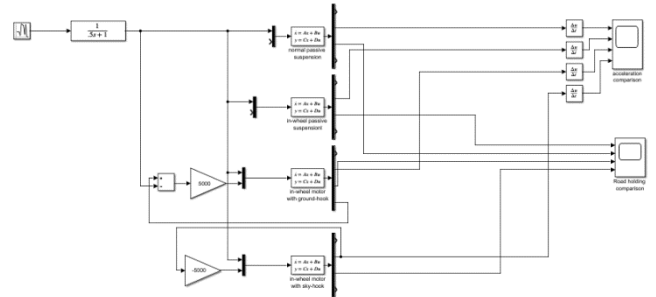


Figure VI. MATLAB simulation for quarter vehicle model with normal passive, in-wheel motor passive suspension system, in-wheel motor with ground-hook controller and in-wheel motor with sky-hook controller

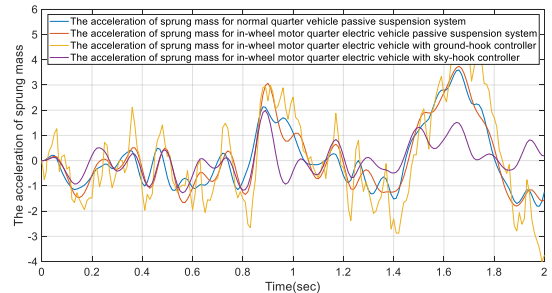


Figure VII. Comparison in acceleration of sprung mass for quarter vehicle between normal passive, IWM EV passive, IWM EV with ground-hook controller and IWM EV with sky-hook controller

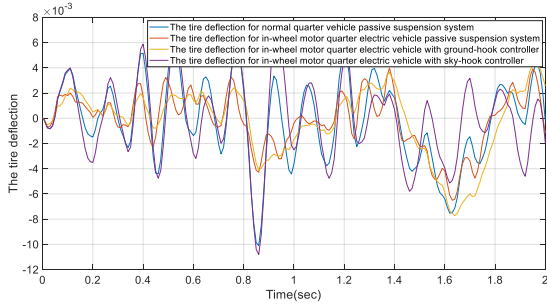


Figure VIII. Comparison in tire deflection for quarter vehicle between normal passive, IWM EV passive, IWM EV with ground-hook controller and IWM EV with sky-hook controller

### III. Half Car Model

#### A. Half Car Model Passive Suspension System

It is presented as a linear four-Degree-of-Freedom (4-DOF) system. The vehicle body has two motions of heave and pitch and the front and rear tires motions are also involved in the half car. A single mass of car is connected to two wheels masses at each corner. Vertical and pitch motion is appropriate for sprung mass while only vertical motion for both unsprung masses. The pitch motion for the half car sprung mass is represented and the vertical displacements for the front tire ( $z_{u1}$ ) and for the rear tire ( $z_{u2}$ ) are also be introduced.

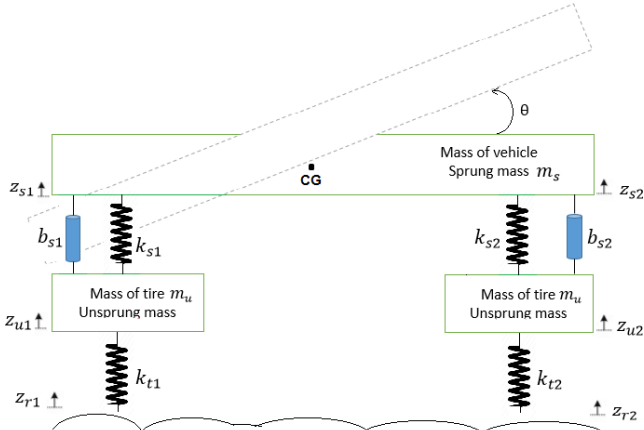


Figure IX. Passive suspension system for half car model (4-DOF)

#### B. Half Car Model Forces

$x$  state vector is  $[z(z_s), v(\dot{z}), (\theta), (\dot{\theta}), z_{u1}, z_{u1}^{\dot{}} , z_{u2}, z_{u2}^{\dot{}}]$   
Position of sprung mass  $z(z_s)$ , velocity of sprung mass  $(\dot{z})$ , the pitch angle  $(\theta)$ , yaw response  $(\dot{\theta})$ ,  $z_{u1}$  (position of front tire),  $z_{u1}^{\dot{}}$  (absolute velocity for unsprung mass front tire),  $z_{u2}$  (position of rear tire),  $z_{u2}^{\dot{}}$  (absolute velocity for unsprung mass rear tire)

$u$  is the input vector  $[z_{r1}, z_{r2}]$ .

$$a_1 = m_s \ddot{z} \quad (27)$$

$$a_2 = k_1 (z_1 - z_{u1}) + b_{s1} (\dot{z}_1 - \dot{z}_{u1}^{\dot{}}) \quad (28)$$

$$a_3 = k_2 (z_2 - z_{u2}) + b_{s2} (\dot{z}_2 - \dot{z}_{u2}^{\dot{}}) \quad (29)$$

$$a_1 + a_2 + a_3 \quad (30)$$

$$z_1 = z - lf \theta \quad (31)$$

$$\dot{z}_1 = \dot{z} - lf \dot{\theta} \quad (32)$$

$$z_2 = z + lr \theta \quad (33)$$

$$\dot{z}_2 = \dot{z} + lr \dot{\theta} \quad (34)$$

$$\ddot{z} = \left( -\frac{k_1}{m_s} - \frac{k_2}{m_s} \right) z + \left( \frac{k_1 lf}{m_s} - \frac{k_2 lr}{m_s} \right) \theta - \left( \frac{b_{s1}}{m_s} + \frac{b_{s2}}{m_s} \right) \dot{z} + \left( \frac{b_{s1} lf}{m_s} - \frac{b_{s2} lr}{m_s} \right) \dot{\theta} + \frac{k_1}{m_s} z_{u1} + \frac{b_{s1}}{m_s} \dot{z}_{u1} + \frac{k_2}{m_s} z_{u2} + \frac{b_{s2}}{m_s} \dot{z}_{u2} \quad (35)$$

$$(m_u + m_i) z_{u1}^{\ddot{}} = -F_t + F_s + F_b \quad (36)$$

$$b_1 = m_t z_{u1}^{\ddot{}} \quad (37)$$

$$b_2 = -k_{t1} (z_{u1} - z_{r1}) + k_1 (z_1 - z_{u1}) \quad (38)$$

$$b_3 = b_{s1} (\dot{z}_1 - \dot{z}_{u1}^{\dot{}}) \quad (39)$$

$$b_1 = b_2 + b_3 \quad (40)$$

$$c_1 = z_{u1}^{\ddot{}} \quad (41)$$

$$c_2 = -\frac{k_{t1}}{m_t} z_{u1} - \frac{k_1}{m_t} z_{u1} - \frac{b_{s1}}{m_t} \dot{z}_{u1}^{\dot{}} \quad (42)$$

$$c_3 = \frac{k_1}{m_t} z + \frac{b_{s1}}{m_t} \dot{z} \quad (43)$$

$$c_4 = -\frac{k_1 lf}{m_t} \theta - \frac{b_{s1} lf}{m_t} \dot{\theta} \quad (44)$$

$$c_5 = \frac{k_{t1}}{m_t} z_{r1} \quad (45)$$

$$c_1 = c_2 + c_3 + c_4 + c_5 \quad (46)$$

$$d_1 = m_t z_{u2}^{\ddot{}} \quad (47)$$

$$d_1 = -f_t + f_s + f_b \quad (48)$$

$$d_2 = -k_{t2} (z_{u2} - z_{r2}) + k_2 (z_2 - z_{u2}) \quad (49)$$

$$d_3 = b_{s2} (\dot{z}_2 - \dot{z}_{u2}^{\dot{}}) \quad (50)$$

$$d_1 = d_2 + d_3 \quad (51)$$

$$e_1 = z_{u2}^{\ddot{}} \quad (52)$$

$$e_2 = -\frac{k_{t2}}{m_t} z_{u2} - \frac{k_2}{m_t} z_{u2} - \frac{b_{s2}}{m_t} \dot{z}_{u2}^{\dot{}} \quad (53)$$

$$e_3 = \frac{k_2}{m_t} z + \frac{b_{s2}}{m_t} \dot{z} \quad (54)$$

$$e_4 = \frac{k_2 lr}{m_t} \theta + \frac{b_{s2} lr}{m_t} \dot{\theta} \quad (55)$$

$$e_5 = \frac{k_{t2}}{m_t} z_{r2} \quad (56)$$

$$e_1 = e_2 + e_3 + e_4 + e_5 \quad (57)$$

$$g_1 = I \ddot{\theta} - k_1 (z_1 - z_{u1}) lf + k_2 (z_2 - z_{u2}) lr \quad (58)$$

$$g_2 = -b_{s1}(\dot{z}_1 - \dot{z}_{u1})lf + b_{s2}(\dot{z}_2 - \dot{z}_{u2})lr \quad (59)$$

$$g_3 = g_1 + g_2 \quad (60)$$

$$r_1 = \left( \frac{k_1 lf}{I} - \frac{k_2 lr}{I} \right) z + \left( \frac{b_{s1} lf}{I} - \frac{b_{s2} lr}{I} \right) \dot{z} \quad (61)$$

$$r_2 = - \left( \frac{k_1 lf^2}{I} + \frac{k_2 lr^2}{I} \right) \theta - \left( \frac{b_{s1} lf^2}{I} + \frac{b_{s2} lr^2}{I} \right) \dot{\theta} \quad (62)$$

$$r_3 = - \frac{k_1 lf}{I} z_{u1} - \frac{b_{s1} lf}{I} \dot{z}_{u1} \quad (63)$$

$$r_4 = \frac{k_2 lr}{I} z_{u2} + \frac{b_{s2} lr}{I} \dot{z}_{u2} \quad (64)$$

$$\ddot{\theta} = r_1 + r_2 + r_3 + r_4 \quad (65)$$

### Parameter for half vehicle model passive suspension system

Mass of car  $m_s=1200$  kg, mass of tire  $m_u=40$  kg, the right spring stiffness  $k_1=16000$ , the left spring stiffness  $k_2=16000$ , the right tire stiffness  $k_{t1}=160000$ , the left tire stiffness  $k_{t2}=160000$ , the right damper  $b_{s1}=1000$ , the left damper  $b_{s2}=1000$ , the distance from the front wheel to the center of gravity ( $lf=1.1$  m), the distance from the rear wheel to the center of gravity ( $lr=1.3$  m),  $I$  is the mass moment of inertia.

$$B = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -\frac{lf}{I} & \frac{lr}{I} \\ 0 & 0 & 0 & 0 \\ \frac{K_{t1}}{m_u} & 0 & 0 & 0 \\ \frac{m_u}{m_u} & 0 & 0 & 0 \\ 0 & \frac{K_{t2}}{m_u} & 0 & 0 \\ 0 & m_u & 0 & 0 \end{bmatrix} \quad (66)$$

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -\frac{K_1}{m_s} - \frac{K_2}{m_s} & -\left( \frac{b_{s1}}{m_s} + \frac{b_{s2}}{m_s} \right) & \frac{K_1 lf}{m_s} - \frac{K_2 lr}{m_s} & \frac{b_{s1} lf}{m_s} - \frac{b_{s2} lr}{m_s} & \frac{K_1}{m_s} & \frac{b_{s1}}{m_s} & \frac{K_2}{m_s} & \frac{b_{s2}}{m_s} \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ \frac{K_1 lf}{I} - \frac{K_2 lr}{I} & \frac{b_{s1} lf}{I} - \frac{b_{s2} lr}{I} & -\left( \frac{K_1 lf^2}{I} + \frac{K_2 lr^2}{I} \right) & -\frac{b_{s1} lf^2}{I} - \frac{b_{s2} lr^2}{I} & -\frac{K_1 lf}{I} & -\frac{b_{s1} lf}{I} & \frac{K_2 lr}{I} & \frac{b_{s2} lr}{I} \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ \frac{K_1}{m_u} & \frac{b_{s1}}{m_u} & \frac{K_1 lf}{m_u} & -\frac{b_{s1} lf}{m_u} & -\frac{K_{t1}}{m_u} - \frac{K_1}{m_u} & -\frac{b_{s1}}{m_u} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ \frac{K_2}{m_u} & \frac{b_{s2}}{m_u} & \frac{K_2 lr}{m_u} & \frac{b_{s2} lr}{m_u} & 0 & 0 & -\frac{K_{t2}}{m_u} - \frac{K_2}{m_u} & -\frac{b_{s2}}{m_u} \end{bmatrix} \quad (67)$$

### Parameter for in-wheel motor half electric vehicle (IWM EV) passive suspension system

Mass of car  $m_s=1200$  kg, mass of tire  $m_u=40$  kg, mass of in-wheel motor  $m_i=45$  kg, the right spring stiffness  $k_1=16000$ , the left spring stiffness  $k_2=16000$ , the right tire stiffness  $k_{t1}=160000$ , the left tire stiffness  $k_{t2}=160000$ , the right damper  $b_{s1}=1000$ , the left damper  $b_{s2}=1000$ , the distance from the front wheel to the center of gravity ( $lf=1.1$  m), the distance from the rear wheel to the center of gravity ( $lr=1.3$  m).

The A matrix is shown in eq. 70.

$$B = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -\frac{lf}{I} & \frac{lr}{I} \\ 0 & 0 & 0 & 0 \\ \frac{K_{t1}}{m_t} & 0 & 0 & 0 \\ \frac{m_t}{m_t} & 0 & 0 & 0 \\ 0 & \frac{K_{t2}}{m_t} & 0 & 0 \\ 0 & m_t & 0 & 0 \end{bmatrix} \quad (68)$$

### Parameter for in-wheel motor half electric vehicle (IWM EV) with ground-hook controller

Mass of car  $m_s=1200$  kg, mass of tire  $m_u=40$  kg, mass of in-wheel motor  $m_i=45$  kg, the right spring stiffness  $k_1=16000$ , the left spring stiffness  $k_2=16000$ , the right tire stiffness  $k_{t1}=160000$ , the left tire stiffness  $k_{t2}=160000$ , the distance from the front wheel to the center of gravity ( $lf=1.1$  m), the distance from the rear wheel to the center of gravity ( $lr=1.3$  m), the ground-hook damping coefficient  $c_{grad1}, c_{grad2}=10000$ ,  $z_{r1}$  (random ground input 1),  $z_{r2}$  (random ground input 2),  $f_{grad1}$  (ground-hook force 1),  $f_{grad2}$  (ground-hook force 2)

$u$  is the input vector  $[z_{r1}, z_{r2}, f_{grad1}, f_{grad2}]$ .

$$B = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{m_s} & \frac{1}{m_s} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -\frac{lf}{I} & \frac{lr}{I} \\ 0 & 0 & 0 & 0 \\ \frac{K_{t1}}{m_t} & 0 & -\frac{1}{m_t} & 0 \\ 0 & 0 & -\frac{1}{m_t} & 0 \\ 0 & \frac{K_{t2}}{m_t} & 0 & -\frac{1}{m_t} \end{bmatrix} \quad (69)$$

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ -\frac{K_1}{m_s} - \frac{K_2}{m_s} & -\left(\frac{b_{s1}}{m_s} + \frac{b_{s2}}{m_s}\right) & \frac{K_1 l f}{m_s} - \frac{K_2 l r}{m_s} & \frac{b_{s1} l f}{m_s} - \frac{b_{s2} l r}{m_s} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{K_1 l f}{I} - \frac{K_2 l r}{I} & \frac{b_{s1} l f}{I} - \frac{b_{s2} l r}{I} & -\left(\frac{K_1 l f^2}{I} + \frac{K_2 l r^2}{I}\right) & -\frac{b_{s1} l f^2}{I} - \frac{b_{s2} l r^2}{I} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{K_1}{m_t} & \frac{b_{s1}}{m_t} & \frac{K_1 l f}{m_t} & -\frac{b_{s1} l f}{m_t} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{K_2}{m_t} & \frac{b_{s2}}{m_t} & \frac{K_2 l r}{m_t} & \frac{b_{s2} l r}{m_t} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad (70)$$

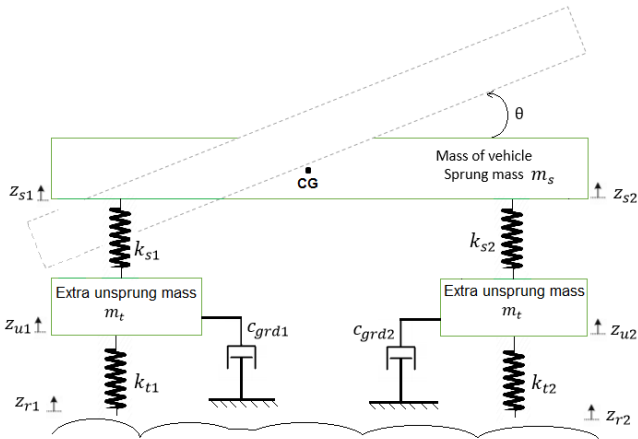


Figure X. In-wheel motor half electric vehicle with ground-hook controller

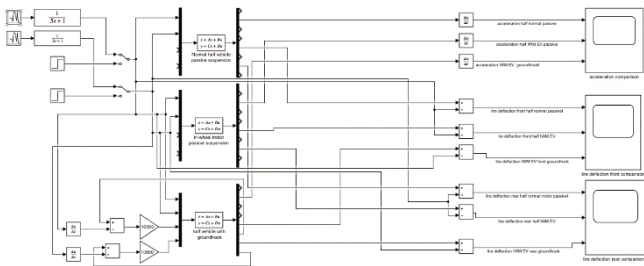


Figure XI. MATLAB simulation comparison for half vehicle between normal passive, in-wheel motor passive and in-wheel motor with ground-hook controller

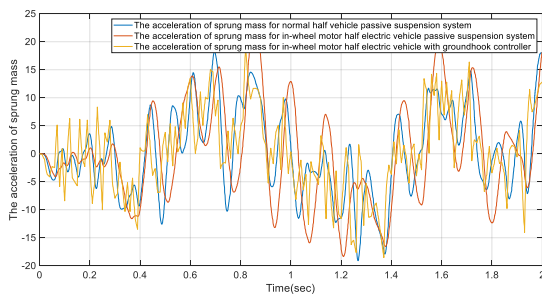


Figure XII. Comparison in acceleration of sprung mass for half vehicle between normal passive in-wheel motor passive and in-wheel motor with ground-hook controller

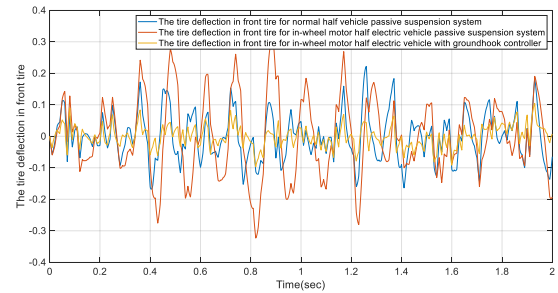


Figure XIII. Comparison in tire deflection front tire for half vehicle between normal passive in-wheel motor passive and in-wheel motor with ground-hook controller

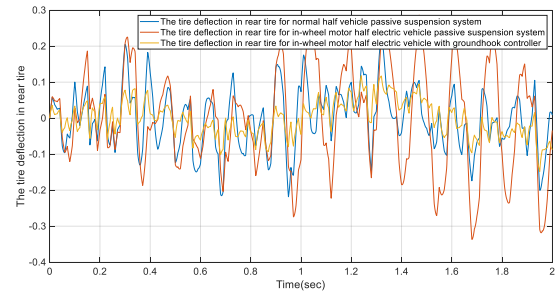


Figure XIV. Comparison in tire deflection rear tire for half vehicle between normal passive, in-wheel motor passive and in-wheel motor with ground-hook controller

### C. Simulation and Analysis

Figures IV, V, VI, VII shows the simulations for the passive and active quarter car models. Figures XII, XIII and XIV shows the simulations for the passive and active half car models. Figure IV, VI and Figure XII compares the accelerations of the vehicle bodies for the normal passive vehicle, the (IWM) passive vehicle, the (IWM) ground-hook active suspension vehicle and the (IWM) sky-hook active suspension vehicle. The acceleration is commonly used to measure passenger comfort. According to Figure IV, VI and Figure XII the added weight of the in-wheel electric motor does not have much of an impact on the acceleration and vehicle comfort. However, the ground-hook active suspension increases the acceleration and diminishes passenger comfort while the sky-hook active suspension decreases the acceleration and improves the comfort for the occupants inside the car. Figure V, Figure VII, Figure XIII and Figure XIV show the tire deflections for the normal passive vehicle, the (IWM) passive vehicle, the (IWM) ground-hook active suspension vehicle and the (IWM) sky-hook active suspension vehicle. The added weight of the in-wheel motor



increases the tire deflection dramatically. Therefore, it is observed that the road holding properties of the tire is worse for the in-wheel motor passive suspension system and also worse for sky-hook controller active suspension because the sky-hook controller only improves the comfort of the passengers. However, the ground-hook controller active suspension is able to decrease the tire deflection to values close to the passive suspension. Therefore, the ground-hook controller is able to reach its goal of improving road holding. This simulation shows that the ground-hook controller is able to eliminate the negative effects of an (IWM EV) in terms of road holding. However, this is done at the cost of decreasing passenger comfort. The sky-hook controller is able to improve the comfort for the passengers of an (IWM EV), and fails to improve the road holding. It should be noted that the gain of the controller for the ground-hook controller can be adjusted to decrease the negative effects on passenger comfort. However, this will decrease the effectiveness of the ground-hook controller.

#### IV. Conclusion

A passive suspension system for both normal passive vehicle and IWM EV without any controller, IWM EV with ground-hook controller and IWM EV with sky-hook controller were analyzed using MATLAB/Simulink. The simulations show that the ground-hook controller improves road holding by reducing the tire deflection for the IWM EV. The sky-hook is shown to improve the comfort for the occupants inside the car. In the case of decreasing or increasing unsprung mass, there is an opposite relationship between size of unsprung mass and road holding. When the unsprung mass is decreased, better road holding response is achieved. On the other hand, increasing the unsprung mass for IWM EV (adding the mass of electric motor 45 kg), by increasing the unsprung mass of the front and rear wheel assemblies results in worse road holding which directly affects the ground contact with the tire. This is the reason for worse road holding for IWM EV. When the ground-hook controller is used for IWM EV, better road holding and good contact with the road is achieved. This is because the ground-hook controller provides better isolation from road disturbances by reducing upper and lower peaks of the tire deflection. But the sky-hook controller fails to improve the road holding because it is used to provide comfort for the passengers. These results were

shown with simulations using both the quarter car model and the half car model.

#### References

- [1] Nouby M. Ghazaly, Ahmad O. Moaaz, "The Future Development and Analysis of Vehicle Active Suspension System", Journal of Mechanical and Civil Engineering (IOSR-JMCE) 2014.
- [2] Vivek Kumar Maurya and Narinder Singh Bhangal, "Optimal Control of Vehicle Active Suspension System", Journal of Automation and Control Engineering 2018.
- [3] John Ekoru "Intelligent Model Predictive/Feedback Linearization Control of Half-Car Vehicle Suspension Systems" thesis 2012.
- [4] Fischer, D. and Isermann, "Mechatronic semi-active and active vehicle suspensions", Control Engineering Practice, 2004.
- [5] Canale, M., Milanese, M. and Novara, "Semi-active suspension control using fast model-predictive techniques", IEEE Transactions on Control Systems Technology, 2006.
- [6] M. FARID ALADDIN, JASDEEP SINGH, "MODELLING AND SIMULATION OF SEMI-ACTIVE SUSPENSION SYSTEM FOR PASSENGER VEHICLE", Journal of Engineering Science and Technology, 2018.
- [7] M. VALÁŠEK , M. NOVÁK , Z. ŠIKA & O. VACULÍN, "Extended Ground-Hook - New Concept of SemiActive Control of Truck's Suspension", Extended GroundHook - New Concept of Semi-Active Control of Truck's Suspension, Vehicle System Dynamics, 2007.
- [8] Hossein Salma ni, Milad Abbasi, Tondar Fahimi Zand , Mohammad Fard and Reza Nakhaie Jazar, "A new criterion for comfort assessment of in-wheel motor electric vehicles", Journal of Vibration and Control, 2020.
- [9] Yechen Qin, Chenchen He, Peng Ding, Mingming Dong, Yanjun Huang, "suspension hybrid control for in-wheel motor driven electric vehicle with dynamic vibration absorbing structures", IFAC Papers OnLine 51-31 (2018).
- [10] Fangwu Ma, Jiawei Wang, Yang Wang, Longfan Yang, 2969. "Optimization design of semi-active controller for in-wheel motors suspension", 2018.
- [11] Jialing Yao, Shuhui Ding, Zhaochun Li, Shan Ren, Saied Taheri, Zhongnan Zhang, "Composite Control and Co-Simulation in In-Wheel Motor Electric Vehicle Suspension", The Open Automation and Control Systems Journal, 2015.
- [12] Patil, I. and Wani, K., "Design and Analysis of Semi-active Suspension Using Skyhook, Ground Hook and Hybrid Control Models for a Four Wheeler", SAE Technical Paper 2015.
- [13] Pipit Wahyu Nugroho, Weihua Li, Haiping Du, Gursel Alici, and Jian Yang, "An Adaptive Neuro Fuzzy Hybrid Control Strategy for a Semiactive Suspension with Magneto Rheological Damper", Hindawi Publishing Corporation Advances in Mechanical Engineering, 2014.
- [14] Rajesh Rajamani, "Vehicle Dynamics and Control", Department of Mechanical Engineering University of Minne.

# Evolutionary Deep automatic CAD system for Early detection of Diabetic Retinopathy and its severity classification.

Lamiya Salman  
Faculty of Engineering & Information  
Science, University of Wollongong in  
Dubai, Dubai, UAE  
lamiyasalman3@gmail.com

Dr Nidhal Abdulaziz  
Faculty of Engineering & Information  
Science, University of Wollongong in  
Dubai, Dubai, UAE  
NidhalAbdulaziz@uowdubai.ac.ae

**Abstract**— Diabetic retinopathy (DR) is a retinal malady prevalent in individuals aged 25 years and above, that leads to irreversible blindness when left unchecked. DR has yet to find a definitive cure. Diagnostic and corrective measures may prevent complete blindness by up to 90% provided early screening and monitory clinical visits. Although ample research and advancements have been made, none have been successfully integrated into the medical system. This is due to the lack of acceptable classification accompanying screening. Further, diluted efforts have been made towards monopolizing early detection i.e., detection of non-proliferate DR a subclass and the earliest discernable form of DR. This study aimed at the early detection of DR through adoption of the medically validated 5 class DR severity grading. The proposed system was a novel hybrid scheme based on past works. Multiple input image modalities and classifier combinations were tested to configure the final DR grading scheme. The system was actualized using fundus images as input which were pre-processed and augmented using green channel extraction, CLAHE and binary masking. Images were then synthesized into discriminatory features using ResNet-50. Final system tier consisted of severity grading through classification using novel CNN based Support Vector Machine (SVM). The utilized ensemble of augmentation and pre-processing module along with ResNet-50 based SVM classifier was a novel contribution not explored in any past works. Proposed pre-processing increased system accuracy by 13.9% and 16% on IDRiD and Kaggle. Overall F1-score, SN, SP and Acc of 0.978 ,0.979, 0.995 and 0.979, was achieved. Incorporation of Artificial Intelligence made proposed system time, cost, and labor efficient which is key in DR screening.

**Keywords**— Diabetic Retinopathy, severity classification, Pre-Processing (PP), Convolutional Neural Network (CNN), Support Vector Machine (SVM).

## I. Introduction

Diabetic retinopathy (DR), a progressive sight-threatening ailment, develops in individuals with prolonged diabetes type 1 or 2, due to presence of high blood sugar levels in their system [1]. DR has a predicted global prevalence of 4.4% by 2030 as per World Health Organization and presents an asymptomatic inceptive stage, long latent phase which when deprived from an early diagnosis may eventually lead to blindness[2]. DR diagnosis and severity analysis today, is provided only by an ophthalmologist or through his evaluation of retinal images/scans. This process is cost, time and labor intensive, especially in case of large disproportionate field expert to patient ratios, often leading to no patient screening [1][3]. These alarming statistics, need for a cost efficient and accessible screening system have motivated copious research and advancement in computer aided diagnostic (CAD) devices but no such work has been able to swap places with a comprehensive eye test [5]. This has been due to inadequate system performance, lack of training data and classification mostly limited to binary severity grades. Lastly, past works

have not placed much importance towards early detection of DR.

This study recognizes the vital importance of early detection and aims at a five-level severity grade output. This need for early detection is reinforced by the fact that progression and possible cure can be achieved only in the early stages of DR as per Early Treatment Diabetic Retinopathy Study (ETDRS) and Diabetic Retinopathy Study (DRS) [6]. Scope of project extends strictly to screening, grading and subsequent preventive measures/guidelines.

## II. Theoretical Background

### A. Data Processing

Input retinal images are captured and sourced in real time conditions such as poor contrast, blurring and non-uniform illumination. Counter acting these implications is essential as system performance is dependent on accurate retinal DR feature localization and subsequent classification based on these discriminant traits.

Histogram Equalisation (HE) is a common method of image contrast enhancement used by Foeady *et.al.* [7]. Although frequently used, HE alters the mean brightness of input image introducing artifacts and intensity saturation therefore is not a suitable choice. Contrast limited Adaptive histogram equalization (CLAHE) is a variant of HE which increases image quality, incorporates uniform intensity equalization and reduces amplified noise thereby takes care of the shortcomings of HE. CLAHE, out performs global enhancement methods like contrast stretching in blood vasculature (BV) enhancement as seen in the work done by M.H. Fadzil, *et.al.* [8], this is further validated through highest peak signal to noise ratio and lowest absolute mean brightness error when compared to other variants such as Adaptive Histogram Equalization [8]. Extraction of a single component of the RGB models can be done as a means of contrast enhancement. Empirical evidence along with entropic comparisons by N. R. Binti, *et.al.*[9] support extraction of the green component due to its capacity to provide the best distinction among DR features/lesions such as Micro Aneurysms (MA), Cotton Wool spots, Exudate (EXs), Hemorrhages (HM) and BV, when compared to red, blue and gray channels. This enables maximum entropy preservation while enhancing contrast. In contradiction to this, Optic disk localization is aided by red channel entropy extraction due to better visual contrast as done by S. Pathan, *et.al.*[10]. Qiao, *et.al.*[11] employed Matched Filters (MF) combined with LoG filters to localize transient EXs, MFs are not suitable as they enhance BV along with the EXs. Difference in lesion size will lead to inaccurate localization.

H. Chen *et.al.* [12] cropped retinal images that had incomplete hemispherical boundaries to form complete spheres. This approach is not ideal as removal of peripheral regions would lead to inaccurate severity grading due to loss of diagnostic lesion areas.

While surveying the appropriate literature, a trend was identified among CNN based architectures where pre-processing was mostly limited to Data Augmentation (DA) such as horizontal, vertical flipping, rotation, mirroring, scaling and grayscale normalization as seen in work done by S. Pao, *et.al.* [13]. This may have been due to the added pre-processing effect of the Max pooling layer by means of de noising along with the prevalence of DA which exposes a variety of variations to the concerned Neural Networks, aiding feature learning instead of separate enhancement of said features. The direct increase in system performance by means of increased PP does not translate in CNN based systems due to their ability to iteratively learn features from input images without requiring feature specific enhancement or enrichments. Processing may become disruptive to said feature learning process due to dampening of distinctive characteristics. Undue PP leads to added computational and time complexity along with added memory constraints on system. Despite this, contrast enhancement and optimal channel extraction have unequivocally facilitated improved system performance being the choice of many CNN and non-CNN users [14].

## B. Feature Segmentation and Extraction

Features form the basis of classification, supplementing the classifier with information necessary to sort between different classes. This may be done through feature specific, Region of Interest (ROI) and CNN based feature abstraction.

The Optic Disk (OD) and BV present in the retina are a potent source of errors, false positive rates (FPR) and misclassification. The minimum-intensity maximum-solidity (MinIMaS) algorithm was invoked by S. Roychowdhury, *et.al.* [15] for means of OD extraction and masking. Results of overall classification system showed 100% sensitivity (SN) but 53.16% specificity (SP), i.e. high FPR.

Thresholding is a common approach adopted for segmentation. In case of BV, wide range of values are present due to difference in vessel width and edge pixels. M.U. Akram, *et.al.* [16] eradicated this limitation by means of multilayered thresholding. Skeletonization of segmented image by means of thinning morphological operation (MO) was done. Although Acc of 94.69% was achieved, high performance metrics on limited dataset are misleading due to exposure to limited sample types.

MAs are sacular capillary dilatations, for efficient MA segmentation, Mateen *et.al.*[17] applied vessel extraction using a novel hybrid approach of enhanced Gaussian Mixture Model (GMM) using Adaptive Learning Rate (ALR) to aid ROI localization for subsequent feature extraction. Incorporation of ALR was the optimal choice as it helps increase learning rate and convergence speed over others such as Expectation Maximization (EM). Post-vessel detection Connected Components Analysis (CCA) along with blob analysis was utilized to differentiate among lesions and healthy retina to enhance BV extraction. CCA is ideal as

it handles differentiation of BV based on colour, proximity, shape, and size. An accuracy of 98% was reported. MA and HM are spherical in shape with a diameter larger than feeding blood vessels. Using this property, H.F. Jaafer, *et.al.* [18] segmented red lesions by applying the flood-filled operation to pixels pertaining to the background and subtracting them from the original image. Incorrect discrimination between circular and linear shapes may be done due to inaccurate threshold selection. A SN of 89.7% was reported by the study.

V. K. Sree and P. S. Rao, [19] carried out EX segmentation by applying Canny edge detection and smoothing via gaussian mask was done. CCA was used to colour code segmented EXs regions and subtract from original image for exudate localization. Reported accuracy was 72.7%. P. Khojasteh, *et.al.*[20] extracted features without added pre-processing or prior segmentation for means of EXs feature extraction using ResNet-50, Discriminative Restricted Boltzmann Machines (DRBMs) and a custom CNN. Feature extraction was enhanced using PCA along with parameter optimization using grid-search approach. Results of ResNet-50 were encouraging (Acc 98.2%, SN 99%.) but that of custom CNN (Acc 70%) and DRBMs (Acc 89%) were poor.

It may be concluded that while hand crafted feature extraction as seen in [18],[15] is efficient, it is complex and static hence would not fare well in extending system to different datasets while maintaining high level of performance. This issue is reinforced due to manual requirements such as threshold selection based on training data seen in [18]. CNN based feature extraction is robust and fully automated hence is dynamic, having an edge over feature specific non-CNN extractions. High accuracies may be accountable to overfitting of training data. To exploit CNNs for feature extraction, ROI localization, balanced and large datasets by means of augmentation as seen in [17][20] may be employed.

## C. Classification

Classification may be seen at lesion level or at image level into DR severity grades. Previous studies showed several binary and multilevel classifications such as DR/No DR, Referral(R)/ Non-referral(NR) DR, and the ETDR 5-level classification adopted in this project.

### 1) Feature based classification

M. H. Ahmad Fadzil, *et.al.*[8] used a non-conventional approach by employing the analysis of the Foveal Avascular zone (FAZ) for DR detection. Classification was limited to a three-level severity grading with SP(>98%), SN(100%) and Acc of (99%). Although the system fared exceptionally, FAZ area overlap were seen and deemed as progressive stages. M.U. Akram, *et.al.* [16] employed PDR classification based on neovascularization detection. A 15x15 window was slid over segmented BV to compute density and energy. Disparity in behavior with normal blood vessel led to PDR classification. Acc of 95.02% was achieved.

A novel two step lesion classification and severity estimation was done by S. Roychowdhury, *et.al.*[15]. Various feature-based classifiers such as k-nearest neighbors (KNN), GMM, SVM and their ensembles were tested to find best suited pick for the separate classification of bright and red

lesions. System achieved high SN ranging from 100-92% but poor SP 48-58% reflecting a high FPR due to bias in algorithm selected for feature set optimization, feature set being key due to usage of feature-based classifiers.

Akram, *et.al.* [21] monopolized M-Medoid classification to develop a hybrid classification model using probabilistic integration of GMM classifiers, an effective multi-class discriminant aided by intuitive modelling due to their resilience towards class overlap which is common in DR classification classes. Severity classification was based on numeric presence of MAs, EXs and HMs. Reported metrics were promising at 99.17% SN, 97.07% SP and 98.52% Acc.

## 2) CNN

The prevalent issue of class imbalance in CNNs is directly addressed by a novel two-step CNN architecture developed by N. Eftekhari, *et.al.* [22] which was claimed to decrease FPR by means of pixel wise classification of MAs. Large FPR was generated reaching up to 8 FP per image at highest reported SN of 77.1%. Training and developing CNNs manually along with pixel wise computation are labor, time intensive and in case of the former error prone.

Shaban, *et.al.* [23], H. Chen, *et.al.* [12] and H. Seyedarabi, *et.al.* [14] undertook DR severity grading using Transfer learning (TL) without prior feature extraction/segmentation. Authors in [23] engineered DR severity grading using TL applied VGG-19 network. No prior PP was utilized. Highest Acc, SN, and SP of 89%, 89%, 95% were reported. TL enhanced CNN by authors in [23] fared better than the custom CNN proposed by N. Eftekhari, *et.al.*[22]. While opting out of segmentation and feature extraction is acceptable due to structural gains of CNNs, PP such as DA and contrast enhancement are crucial for system performance especially due to the ambiguous nature of retinal features. Classification was limited to No/Moderate/Severe DR. Unlike authors in [23], authors in [14] applied CLAHE for PP and claimed EfficientNet to be the ideal choice of CNN due to their reduced parameters, flops, increased speed and accuracy. Classification was limited to referral and non-referral DR with reported 93% SN. H. Chen, *et.al.* [12] aimed at 5 level severity classification. To this end, pre trained Inception V3 due to its depth and elevated linear expression was enhanced using Stochastic Gradient Decent (SGD) and early stop mechanism to mitigate learning rate and overfitting. Stated F1 score of 77% was poor, especially in class 1 and 3.

## 3) Ensembles

Ensembles of classifiers is an active area of research due to their ability to boost system performance but at the cost of increased system complexity and time constraint [24].

J. Sahlsten, *et.al.* [24] validated this claim by employing the Inception-V3 architecture and its ensemble of 6 units to compare classification performance of the two systems. Results showed a clear increase in performance across all planes when comparing the single CNN against the ensemble system. M. Ghazal, *et.al.*[25] utilized TL and ensembles and combined them to authenticate the advantages of TL. The ensemble of 7 CNNs had pre-trained AlexNets and randomly initialized custom CNNs. Classification was by means of SVM. This choice was made by sampling other classifiers

such as KNN, random forest and random tree. Superiority of TL over custom CNNs was validated by demarcation in performance metrics. This study established the dominance of SVMs over other classifiers when paired with CNNs, through benefits in terms of training, data requirement and time constraint using TL, despite these crucial assertions, system limited itself to early PDR detection.

Qummar, *et.al.*[26] exploiting an ensemble of CNNs, with Resnet50, Inceptionv3, Xception, Dense121, Dense169 as the base models, aimed at encoding rich feature extraction and accurate DR classification into 5 class severities which is not seen often. System output was the stacked combination of the individual predicted class labels. Khalifa, *et.al.* [27] set a comparative study to find the best CNN architecture for the same purpose as authors in [26] but unlike the former was limited to a single CNN. Authors in [26] reported an Acc and F1 score of 80.8% and 53.7% but an impressive F1-score of 95.82% was reported by authors in [27] with the presumption of AlexNet being the best choice and VGG16 being second. CNN classification although robust on its own may be further elevated by interweaving other supervised or un-supervised classifiers forming an ensemble of sorts[25].

P Khojasteh, *et.al.*[20] set up a comparative analysis of a custom CNN, unsupervised DRBM, and ResNet-50 transfer model with its SoftMax classifier layer switched with SVM, supervised Optimum-Path Forest, and k-NN classifiers comparing the performance of each in EX classification. ResNet-50+SVM outperformed others with 0.96 SP,0.99 SN. Mateen *et.al.*[17] classified 5 stages of DR using feature vectors segmented from fundus image utilizing the SoftMax algorithm. Being one of the few 5 grade classifications using a CNN structure reported Acc was of 98.13 %. It is noted that machine learning (ML) classifiers such as GMMs, SVMs are high performing and efficient but are feature set dependent. Hence more complex and resistant to expansion to new datasets[14]. CNNs directly address this issue, and are high performing but are victim to overfitting and class imbalance, combining the two approaches, SVM and CNNs would result in elevated results and resilient classification [17][20].

## III. Materials and Method

The proposed system is a novel hybrid approach designed to generate DR severity grading based on input fundus images. Overall system block diagram is shown in Figure I.

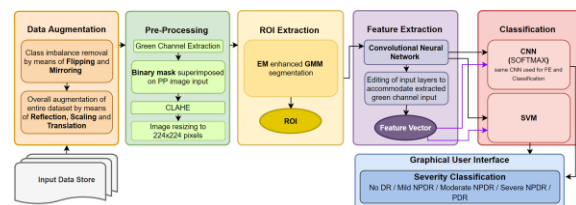


Figure I. Overall System block diagram with all modalities tested.

### 1) Dataset

The KAGGLE benchmark dataset is a compilation of 35,126 images with separate training and testing sets with class labels ranging from 0-4, in line with the severity scheme adopted by this study.

IDRiD comprises of 516 images. Images are available in .jpg format, 4288x2848 resolution with 800KB size. Images are accompanied with severity grading and ground truths.

Table I. Effect of Pre-processing datasets on performance.

Data set	Classifier	No of images	Accuracy	Sensitivity	Specificity	Precision	Recall	F1 Score
IDRiD (raw)	SoftMax	840	0.91	0.760	0.945	0.771	0.767	0.763
IDRiD (PP)	SoftMax	840	0.933	0.824	0.96	0.826	0.824	0.824
Kaggle (raw)	SoftMax	1000	0.836	0.584	0.899	0.59	0.584	0.578
Kaggle (PP)	SoftMax	1000	0.904	0.776	0.944	0.755	0.756	0.743
IDRiD (raw)	SVM	840	0.84	0.64	0.906	0.60	0.64	0.564
IDRiD (PP)	<b>SVM</b>	<b>840</b>	<b>0.979</b>	<b>0.979</b>	<b>0.995</b>	<b>0.978</b>	<b>0.979</b>	<b>0.978</b>
Kaggle (raw)	SVM	1000	0.824	0.54	0.897	0.56	0.54	0.52
Kaggle (PP)	<b>SVM</b>	<b>1000</b>	<b>0.984</b>	<b>0.961</b>	<b>0.99</b>	<b>0.956</b>	<b>0.961</b>	<b>0.96</b>

## 2) Data Augmentation

Of the two datasets employed, IDRiD was first balanced in terms of images per DR severity class using manual up sampling. Once balanced, datasets were further augmented by means of reflection, scaling and translation. This was done to diminish overfitting to a singular class/dataset and provide a larger dataset for TT purposes from the smaller source dataset. DA was done uniformly across all classes of data, maintaining equal ratios to avoid class imbalance.

## 3) Pre-processing

Acquired images had their green channel extracted prior to masking, highlighting foreground from background [7], providing the best distinction between retinal features in the foreground and darker background pixels [9]. Input retinal fundus images were resized to standard input size of ResNet-50.

## 4) ROI localization

BV segmentation was done in accordance to the work presented by M. Mateen, *et.al.* [17]. Discriminant information present in the BV regions was localized and fed to the CNN for feature extraction. GMM was optimized using EM instead of ALR due to comparable performance and limited works on fairly novel ALR implementation with GMMs [21].

## 5) Feature Extraction and Classification

Feature extraction by means of CNNs was opted for due to their ability to automatically learn features instead of requiring a set of hand-crafted features which are time consuming, complex, and fixed. This project aims at the comparative analysis of different classification approaches. The novelty of the proposed system is the usage of ResNet-50 with SVM for DR severity classification which has not been explored earlier. Classification has been implemented in accordance with the work done by P. Khojasteh, *et.al.* [20], hence is an extension of his work from HE classification to DR grading. Multiple classification approaches, ResNet-50 architecture with raw fundus images, pre-processed images, segmented images and ResNet-50 with SVM classifier and raw/ PP/ segmented images have been tested. SVM and ResNet-50 have been selected due to their consistent high performance. Classification without SVM classifier was done by means of SoftMax activation technique, both protocols



Figure II. Kaggle Dataset (no pre-processing): No DR, Mild NPDR, Mod NPDR, Severe NPDR and PDR (L-R).

were based on features extracted by the CNN. 5-fold cross validation was used to reflect authenticity of acquired results certifying lack of overfitting due to the 5 separate folds of TT data used. The SGD which helps minimize cross entropy loss, increase computational efficiency and accelerate learning was used [23][3].

## IV. Results and Discussion

Comparative analysis between Raw, Pre-processed, ROI input modalities and the SVM and SoftMax classifier using the ResNet-50 architecture was done. Multiple mini batch sizes, epochs and learning rates were experimented with to identify the best performing fusion. Through experimentation the optimum combination of 10 mini batch size, 50 epochs and 1e-3 learning rate was ascertained. System performance markedly increased across all parameters on increasing TT images. This outcome was expected as CNN performance is directly proportional to the mass of images supplied to the architecture. More images make way for better learning and feature extraction. It may be inferred from Table I that system accuracy was increase by 13.9 to 16 % by pre-processing input images prior to classification.

Further it may be noted that the smaller IDRiD (840) dataset achieved higher SN and F1 score as compared to the larger Kaggle(1000) dataset. This may be accounted for by the unruly nature of the latter. Kaggle comprises of 35,000 plus images captured by varying fundus cameras, zooming scales, illumination, and FOV angles, Figure II [17]. Inhomogeneous quality and sub-optimal lighting as compared to the uniform composition of the IDRiD dataset, Figure III, explains the higher system performance linked with the latter despite its smaller size. This finding reinforces the importance dataset quality and composition has on performance and eventual incorporation of a computer aided diagnostic DR grading system into the medical field. Effect of Segmentation was tested using only the Kaggle dataset. Overall system performance of the SoftMax classifier was lower than the system performance achieved using only PP images at Acc= 0.89 and F1-score of 0.72. Although performance was expected to improve as per the study executed by M. Mateen, *et.al.* [17], replacement of EM as optimizer in place of the originally proposed ALR may be accounted for the diminished performance. ALR as opposed to EM has an adaptive variable which updates the gaussian clusters through the weight factor parametric elimination to

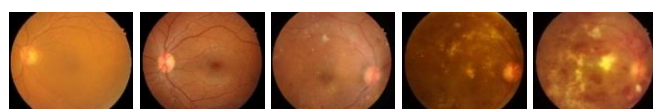


Figure III. IDRiD Dataset (no pre-processing): No DR, Mild NPDR, Mod NPDR, Severe NPDR and PDR (L-R).

Table II. Comparison of proposed method with past works.

Author	Methodology	Data set	Acc	SN	SP	Precision	Recall	F1 Score
M. Shaban, [33]	Custom CNN	Kaggle	0.88	0.87	0.94	-	-	-
N. Khalifa,[35]	Augmentation with AlexNet	Kaggle	0.979	-	-	0.962	0.954	95.82
S. Poa, [19]	Bichannel CNN,	Kaggle	0.878	0.778	0.938	-	-	-
Chen, [18]	Inception V3	Kaggle	-	-	-	0.76	0.80	0.77
P. Khojasteh[28]	ResNet-50 + SVM	DIARETDB1	0.982	0.99	0.96	-	-	-
M. Mateen,[25]	GMM ROI extraction,VGG-19	Kaggle	0.983	-	-	-	-	-
Wang,[32]	Inception V3	Kaggle	0.632	-	-	-	-	-
Qummar, [34]	Ensemble of ResNet-50, Inception V3, Xception, Dense 121 and Dense 169	Kaggle	-	-	0.851	0.634	0.65	0.322
<b>Proposed</b>	<b>ResNet-50 + SVM</b>	<b>IDRiD</b>	<b>.979</b>	<b>0.979</b>	<b>0.995</b>	<b>0.978</b>	<b>0.979</b>	<b>0.978</b>
<b>Proposed</b>	<b>ResNet-50 + SVM</b>	<b>Kaggle</b>	<b>0.984</b>	<b>0.961</b>	<b>0.99</b>	<b>0.956</b>	<b>0.961</b>	<b>0.96</b>

enable quasi linear adaptation. Similarly, the SVM classifier using the ROI extracted images fared very poorly. System Acc peaked at 82% with F1 score of 52.2%. Segmentation of input images significantly reduced TT time as compared to raw or PP images. Although DR detection and diagnosis is not time sensitive i.e., does not need to be done in real time, training of CNNs using PP images may take up to 8 hrs for 1000 images. TT over larger datasets will take longer and may prove troublesome. Hence, segmentation optimization should be done as part of future work to enhance system performance using ROI input using ALR. This will make way for smoother integration of the system in hospitals and care giving facilities. It was inferred that system performance had a linear increase with PP and increase of input images. A 13.9 to 16 % increase in system Acc was achieved using PP. The best performing duo was identified as the ResNet-50 architecture with the SVM classifier using only PP images, seen in Table I. Classification using ResNet-50 with the SoftMax classifier on PP images peaked at an Acc of 93.3% with IDRiD and 90.4% with Kaggle as opposed to 97.9% and 98.4 % respectively using the SVM classifier. Segmentation led to degradation of system performance. This unanticipated dip in performance using ROI localization was due to the inability of the EM algorithm to dynamically segment images like the originally proposed ALR algorithm.

System evaluation has been done by means of parametric comparison of operation indicators such as Acc, SN, SP, Recall, precision, and F1 score. All past research considered for comparison accomplished DR grading using Kaggle bar one of the literatures followed. In comparison to past works, the proposed system fared better than most, Table II. The reported precision by N. Khalifa, et.al.[27] was 0.6% higher than the proposed system. F1 score, recall and Acc was lower. This study takes a novel hybrid approach by building on the work done by M. Mateen, et.al. [17] and P. Khojasteh et.al.[20]. Although P. Khojasteh [20] employed different datasets and was limited to HE detection, Table II presents the acquired performance metrics for purpose of comparison. It may be seen that the proposed hybrid method led to enhanced operation compared to M. Mateen, et.al[17]. In case of P. Khojasteh,[20], SP metrics obtained by the proposed system were better than both reported specificities by 3 to 4.5%. Acc achieved on the DIARETDB1 dataset is improved by the proposed system on Kaggle but falls short by .3 % on the IDRiD dataset. Acc on E-Optha is improved on both datasets utilized by the proposed system. Both SNs reported by authors of [20] are higher than the proposed system by 0.1-2.9 %. Neither of the studies provided F1 scores hence direct

comparison of overall performance was not done. Although the study done by P. Khojasteh, et.al. reported higher SN in both datasets and higher Acc in one dataset, It was limited to HE classification and did not do DR classification or severity grading.The proposed system fared better overall against nearly all past works surveyed. Hence, the system has been successful in improving DR classification and detection performance by incorporating past works and generating an evolutionary deep CAD system.

## V. Conclusion

This study was formulated to tackle the rampant issue of screening and grading Diabetic Retinopathy in large populations due to factors such as cost, scarcity of trained physicians or accessibility. Based on qualitative and quantitative analysis of previous works, A novel hybrid automated deep learning algorithm capable of providing screening, grading and preventive measures has been proposed. Due to its end-to-end nature, the system eradicates the need of trained Ophthalmologists there by alleviating time and workload, seamlessly mitigating human error/bias, restricted access to diagnosis and its high cost in the process. The adopted solution to the research question employed image processing and was developed using MATLAB. Multiple types of input images and classifiers were tested to find the superior performing system. The final system initiated with a rigorous pre-processing and augmentation module followed by feature extraction using ResNet-50. Images were then classified using an SVM classifier. System construction was based on research and experimentation which yielded supremacy of the CNN used. The contribution of this study was the robust pre-processing and augmentation module along with classification of DR using ResNet-50 and SVM which has not been explored in any previous studies. This choice was attributed to the enhanced overall performance necessary for adequate DR classification despite the increased complexity constraints. The system was trained on both IDRiD and KAGGLE datasets using 1000 and 840 images, respectively. A GUI was designed to simulate incorporation of system into health facilities. Simulation results inferred an F1-score, SN, SP and Acc of 0.978 ,0.979, 0.995 and 0.979, respectively. A 13.9 % and 16 % increase in system Acc was achieved on IDRiD and Kaggle respectively when proposed pre-processing and augmentation was used. Acc and SP of literature followed was improved by 0.1-0.8% and 3-4.5% respectively. The final system led to increased performance across all parameters in comparison to almost all past works reviewed. Although best results were anticipated using segmented images, this was not the case due

to inadequacy of employed EM in comparison to originally proposed ALR. Further work may be done to test system performance using ALR in ROI segmentation and PCA/SVD algorithms for the enrichment of feature extraction. Robust feature extraction would supplement system generalizability to new datasets and images, this being crucial for successful system integration into a medical center. Editing the input layer size of the CNN structure as opposed to resizing input images would be a promising experiment. This may be done as a solution to the degraded image quality attained when image resizing is done to fit CNN input frame size. Performance may be optimized using Twin SVMs due to their supremacy over traditional SVMs. As established inter-dataset variances are present, using multiple different datasets would help system adjust to different FOV angles, illumination etc. Finally, results showed system performance to be dependent upon quality of images. Hence, the need for better dataset compilation strategies and guidelines is essential to the development of a robust DR detection and grading computer aided system.

### References

- [1] American Optometric Association, "Diabetic retinopathy." <https://www.aoa.org/healthy-eyes/eye-and-vision-conditions/diabetic-retinopathy?sso=y>
- [2] S. Wild, G. Roglic, A. Green, R. Sicree, and H. King, "Global prevalence of diabetes: estimates for the year 2000 and projections for 2030," *Diabetes Care*, vol. 27, no. 5, pp. 1047–1053, May 2004, doi: 10.2337/diacare.27.5.1047.
- [3] A. A. Alghadyan, "Diabetic retinopathy – An update," *Saudi J. Ophthalmol.*, vol. 25, no. 2, pp. 99–111, Apr. 2011, doi: 10.1016/j.sjopt.2011.01.009.
- [4] M. Mateen, J. Wen, M. Hassan, N. Nasrullah, S. Sun, and S. Hayat, "Automatic Detection of Diabetic Retinopathy: A Review on Datasets, Methods and Evaluation Metrics," *IEEE Access*, vol. 8, pp. 48784–48811, 2020, doi: 10.1109/ACCESS.2020.2980055.
- [5] "Diabetic Retinopathy: A Position Statement by the American Diabetes Association | Diabetes Care." <https://care.diabetesjournals.org/content/40/3/412>
- [6] "Early photocoagulation for diabetic retinopathy. ETDRS report number 9. Early Treatment Diabetic Retinopathy Study Research Group," *Ophthalmology*, vol. 98, no. 5 Suppl, pp. 766–785, May 1991.
- [7] A. Z. Foeady, D. C. R. Novitasari, A. H. Asyhar, and M. Firmansjah, "Automated Diagnosis System of Diabetic Retinopathy Using GLCM Method and SVM Classifier," in *2018 5th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI)*, Oct. 2018, pp. 154–160. doi: 10.1109/EECSI.2018.8752726.
- [8] M. H. Ahmad Fadzil, Nor Fariza Ngah, T. M. George, L. I. Izhar, H. Nugroho, and H. A. Nugroho, "Analysis of foveal avascular zone in colour fundus images for grading of diabetic retinopathy severity," in *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology*, Aug. 2010, pp. 5632–5635. doi: 10.1109/IEMBS.2010.5628041.
- [9] N. R. Binti Sabri and H. B. Yazid, "Image Enhancement Methods For Fundus Retina Images," in *2018 IEEE Student Conference on Research and Development (SCORED)*, Nov. 2018, pp. 1–6. doi: 10.1109/SCORED.2018.8711106.
- [10] Sumaiya Pathan, P. Kumar, R. Pai, and S. V. Bhandary, "Automated detection of optic disc contours in fundus images using decision tree classifier," *Biocybern. Biomed. Eng.*, vol. 40, no. 1, pp. 52–64, Jan. 2020, doi: 10.1016/j.bbe.2019.11.003.
- [11] L. Qiao, Y. Zhu, and H. Zhou, "Diabetic Retinopathy Detection Using Prognosis of Microaneurysm and Early Diagnosis System for Non-Proliferative Diabetic Retinopathy Based on Deep Learning Algorithms," *IEEE Access*, vol. 8, pp. 104292–104302, 2020, doi: 10.1109/ACCESS.2020.2993937.
- [12] H. Chen, X. Zeng, Y. Luo, and W. Ye, "Detection of Diabetic Retinopathy using Deep Neural Network," in *2018 IEEE 23rd International Conference on Digital Signal Processing (DSP)*, Nov. 2018, pp. 1–5. doi: 10.1109/ICDSP.2018.8631882.
- [13] Shu-I. Pao, H.-Z. Lin, K.-H. Chien, M.-C. Tai, J.-T. Chen, and G.-M. Lin, "Detection of Diabetic Retinopathy Using Bichannel Convolutional Neural Network," *Journal of Ophthalmology*, Jun. 20, 2020. <https://www.hindawi.com/journals/joph/2020/9139713/>
- [14] H. Seyedarabi, S. H. A. Jahromi, A. Javadzadeh, and ASRA MOMENI POUR, "Automatic Detection and Monitoring of Diabetic Retinopathy Using Efficient Convolutional Neural Networks and Contrast Limited Adaptive Histogram Equalization," *IEEE Access*, vol. 8, pp. 136668–136673, 2020, doi: 10.1109/ACCESS.2020.3005044.
- [15] S. Roychowdhury, D. D. Koozekanani, and K. K. Parhi, "DREAM: Diabetic Retinopathy Analysis Using Machine Learning," *IEEE J. Biomed. Health Inform.*, vol. 18, no. 5, pp. 1717–1728, Sep. 2014, doi: 10.1109/JBHI.2013.2294635.
- [16] M. U. Akram, I. Jamal, A. Tariq, and J. Imtiaz, "Automated segmentation of blood vessels for detection of proliferative diabetic retinopathy," in *Proceedings of 2012 IEEE-EMBS International Conference on Biomedical and Health Informatics*, Jan. 2012, pp. 232–235.
- [17] Muhammad Mateen, J. Wen, Nasrullah, S. Song, and Z. Huang, "Fundus Image Classification Using VGG-19 Architecture with PCA and SVD," *Symmetry*, vol. 11, no. 1, Art. no. 1, Jan. 2019, doi: 10.3390/sym11010001.
- [18] H. F. Jaafar, A. K. Nandi, and W. Al-Nuaimy, "Automated detection of red lesions from digital colour fundus photographs," in *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Aug. 2011, pp. 6232–6235. doi: 10.1109/IEMBS.2011.6091539.
- [19] V. K. Sree and P. S. Rao, "Diagnosis of ophthalmologic disorders in retinal fundus images," in *The Fifth International Conference on the Applications of Digital Information and Web Technologies (ICADIWT 2014)*, Feb. 2014, pp. 131–136. doi: 10.1109/ICADIWT.2014.6814696.
- [20] Parham Khojasteh *et al.*, "Exudate detection in fundus images using deeply-learnable features," *Comput. Biol. Med.*, vol. 104, pp. 62–69, Jan. 2019, doi: 10.1016/j.combiomed.2018.10.031
- [21] M. Usman Akram, S. Khalid, A. Tariq, S. A. Khan, and F. Azam, "Detection and classification of retinal lesions for grading of diabetic retinopathy," *Comput. Biol. Med.*, vol. 45, pp. 161–171, Feb. 2014, doi: 10.1016/j.combiomed.2013.11.014.
- [22] Noushin Eftekhari, H.-R. Pourreza, M. Masoudi, K. Ghiasi-Shirazi, and E. Saeedi, "Microaneurysm detection in fundus images using a two-step convolutional neural network," *Biomed. Eng. OnLine*, vol. 18, no. 1, p. 67, May 2019, doi: 10.1186/s12938-019-0675-9.
- [23] M. Shaban *et al.*, "A convolutional neural network for the screening and staging of diabetic retinopathy," *PLoS ONE*, vol. 15, no. 6, Jun. 2020, doi: 10.1371/journal.pone.0233514.
- [24] Jaakko Sahlsten *et al.*, "Deep Learning Fundus Image Analysis for Diabetic Retinopathy and Macular Edema Grading," *Sci. Rep.*, vol. 9, no. 1, Art. no. 1, Jul. 2019, doi: 10.1038/s41598-019-47181-w.
- [25] M. Ghazal, S. S. Ali, A. H. Mahmoud, A. M. Shalaby, and A. El-Baz, "Accurate Detection of Non-Proliferative Diabetic Retinopathy in Optical Coherence Tomography Images Using Convolutional Neural Networks," *IEEE Access*, vol. 8, pp. 34387–34397, 2020, doi: 10.1109/ACCESS.2020.2974158.
- [26] S. Qummar *et al.*, "A Deep Learning Ensemble Approach for Diabetic Retinopathy Detection," *IEEE Access*, vol. 7, pp. 150530–150539, 2019, doi: 10.1109/ACCESS.2019.2947484.
- [27] Nour Eldeen M. Khalifa, M. Loey, M. H. N. Taha, and H. N. E. T. Mohamed, "Deep Transfer Learning Models for Medical Diabetic Retinopathy Detection," *Acta Inform. Medica*, vol. 27, no. 5, pp. 327–332, Dec. 2019, doi: 10.5455/aim.2019.27.327-332.

# Studies on Increasing Energy Efficiency By Modernization of a Single-Stage Type Turbo-Compressor for Ammonia Combustion Air Process

Oguzhan Erbas  
Department of Mechanical Engineering  
Kutahya Dumlupinar University  
Kutahya, Turkey  
oguzhan.eras@dpi.edu.tr

Ahmet Akbulut  
Department of R&D  
Istanbul Gubre Sanayi A.S (IGSAS)  
Kutahya, Turkey  
ahmet.akbulut@igsas.com.tr

Fadime Menekse Ikbal  
Department of Mechanical Engineering  
Kutahya Dumlupinar University  
Kutahya, Turkey  
menekseikbal@hotmail.com

**Abstract**— Today, due to global warming, security of supply, and rising prices, energy efficiency studies come to the fore. Energy efficiency also seems to be a way for industrial enterprises to be profitable. This study discussed the modernization of a turbo-compressor located within a dilute nitric acid production plant. This single-stage turbo-compressor sends ammonia combustion air into the reactor. As a result of the modernization, the energy efficiency in the system was analyzed. As a result of modernization studies, the speed of the turbo compressor and the amount of air has been increased. Thus, the acid production amount was raised in the facility.

**Keywords**— turbo compressor, modernization, nitric acid production, energy efficiency

## I. Introduction

Energy efficiency is a set of interdisciplinary strategic activities that complement and support national strategic objectives. Significantly reducing the burden of energy costs, ensuring supply security in energy, reducing foreign dependency, applying low-carbon technologies, protecting the environment, using domestic energy potential, and ensuring its sustainability are the primary targets. In addition, it is possible to define energy efficiency as providing the same production capacity using less energy. In other words, energy efficiency is the reduction of the amount of energy consumed without affecting economic development and social welfare without reducing the amount of output and quality in production. Therefore, savings are the most important factors that stand out in terms of energy efficiency. Here, saving intends to minimize energy consumption without hindering economic development and standard of living.

To make progress in energy efficiency, the industrial sector must be considered as a priority. When examining the countries that stand out in energy efficiency worldwide, their improvement in the industrial sector is noticeable. The most intensive energy consumption sectors in Turkey are industry, electricity generation, transportation, and housing. Therefore, these sectors also stand out in terms of energy efficiency. When the sectors are evaluated separately, it is seen that a large number of regulations related to energy efficiency have been made in Turkey, especially in the industrial field. Many energy conversion systems are used in the production, transmission, and storage of energy. Because compressed air is convenient and safe, it is widely used as a power source in control valves, air motors, air guns for cleaning purposes, and many more. Compressed air systems have a low power-to-weight ratio and a high power density. Being resistant to explosions and overloads and not being affected by

temperature, humidity, dust, electromagnetic noise are essential features. In addition, they are preferred by many businesses because they are easy to maintain and can be carried over a long distance[1,2].

Turbo compressors are dynamic compressors that are commonly used to press air and gas. These machines create pressure according to the dynamic principle; this means that pressure increase is provided (using the air velocity) without any mechanical volume contraction (sliding) as in the operation of positive displacement compressors. In turbo compressors, the element that rotates at high speed to push the air (gas) is called an impeller or turbofan. There is no piston or other type of mechanical driving or compression element between the air inlet and the air outlet of the turbocharger[3,4]. Instead, the turbo compressor sucks the air from the suction port (middle), and the impeller blades rotating at high speed create a centrifugal force and blow the air from the inside out (around). Therefore, turbo compressors are also called centrifugal or even aerodynamic compressors. A radial discharge flow characterizes centrifugal compressors. Air is sucked towards the center of a rotating impeller (turbine) with radial blades and is pressed against the circumference of the impeller by centrifugal (centripetal) forces. This circumferential (radial) movement of air causes both pressure increase and kinetic energy generation. Before the air is directed to the turbine center of the following compression stage, it passes through a diffuser and spiral, during which kinetic energy is converted into pressure. Efficiency analysis is critical because these systems are energy-intensive systems [5,6]. Modernization is essential for the continuity of production and energy efficiency today because modernization enables investments to be made in the production lines of existing facilities, including adding suitable parts to machinery and equipment that have completed their technical and/or economic life or replacing existing machinery and equipment with new ones, completing missing parts in the facility, directly raising the quality of the final product or changing its model.

In this study, the modernization of the turbo-compressor system in a dilute nitric acid production facility is discussed. To produce acid in this facility, it is necessary to obtain a mixture of ammonia and air as required by the process. Therefore, the plant capacity is directly dependent on the amount of air produced by this turbo-compressor. Turbo air compressor is operated by a steam turbine and auxiliary gas expansion turbine. With the modernization works; It is aimed to increase the amount of acid production by increasing the speed of the turbo-compressor and the amount of air.



## II. The catalytic ammonia oxidation process

In nitric acid production, first ammonia is oxidized to produce nitric oxide (nitrogen monoxide, NO), then NO is oxidized, one more step to becoming NO<sub>2</sub> (nitrogen dioxide). In the final stage, NO<sub>2</sub> gas is absorbed with water, and nitric acid (HNO<sub>3</sub>) is obtained. In the basic production technology

known as the Ostwald process, the reactions between ammonia and oxygen. For example, oxidation, etc. Further nitrogen oxides are formed between the steps, depending on the pressure and temperature conditions. Many unit processes and processes with catalysts are applied to create this simple reaction chain based on facilities. The general flow chart of the nitric acid production process is shown in Figure I.

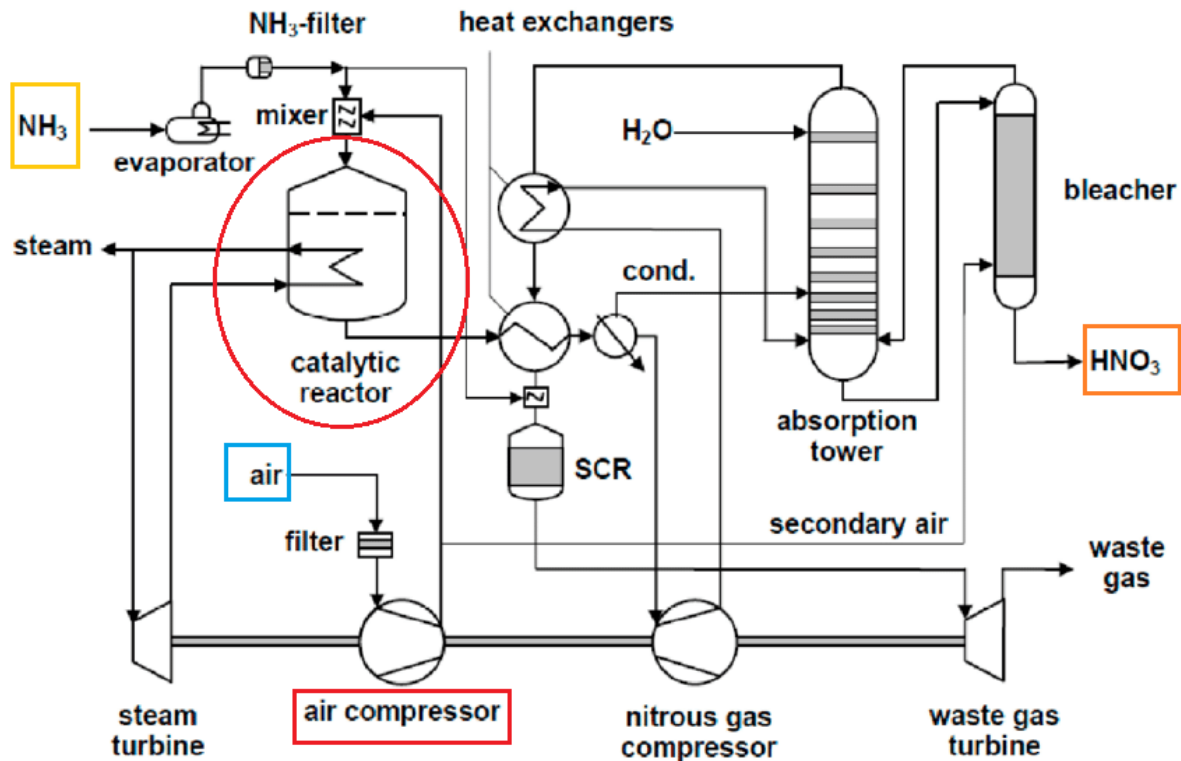


Figure I. Nitric acid production process general flow chart

Nitric acid plants are classified according to different or same pressure levels in the reactors in two separate oxidation stages. Some technologies have the same pressure in both stages; these are single-pressure systems. Others operate at two different pressures, and these are dual pressure systems.

Whether the pressure levels used are low, medium or high pressure also significantly affects the result. The most commonly used are medium and high-pressure dual systems. High-pressure single systems are also widely used. This classification is essential in terms of emissions. The nitric acid production steps are shown in Figure II.

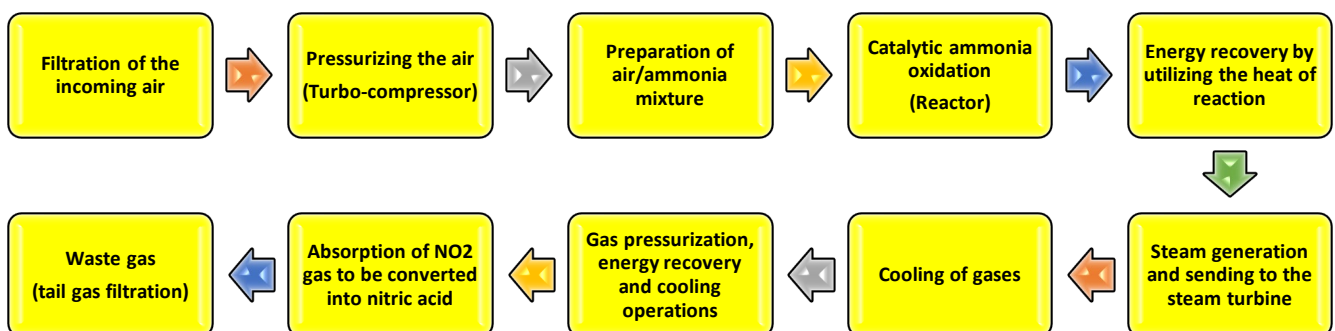


Figure II. The nitric acid production steps

### III. Results and recommendations

The raw materials used in the production of nitric acid are water, air, and ammonia. Ammonia is first gasified with water in the ammonia gasifier and then comes to the ammonia superheaters to be heated with hot air. Then the heated gaseous ammonia comes to ammonia incinerators. The air required for the combustion of ammonia is sucked from the atmosphere by the “turbo-compressor”. The turbo-compressor is operated by the steam turbine and by the rest gas expansion turbine.

Air leaks are one of the most important causes of unnecessary energy consumption in a compressed air system. Necessary preliminary leak studies (leak detection and repair) and continuous system performance reviews ensure that almost all air production in the compressed air

system is converted into useful work. As a result, system energy consumption is realized only in line with operational needs. In addition, since the drops that may occur in the system outlet pressure due to leaks are prevented, inefficient operation of the equipment at the end-use points is controlled, and the continuation of the production processes is ensured [7,8]. First of all, to increase the efficiency of the turbo-compressor operating in the facility, modernization of the sealing system of the rotor is envisaged. For this purpose, new fixed sealing parts were manufactured. The before and after view of the modernized sealing system of the turbo-compressor is shown in Figure III. In addition, the manufacturing drawings of the newly designed sealing elements are shown in Figure IV.

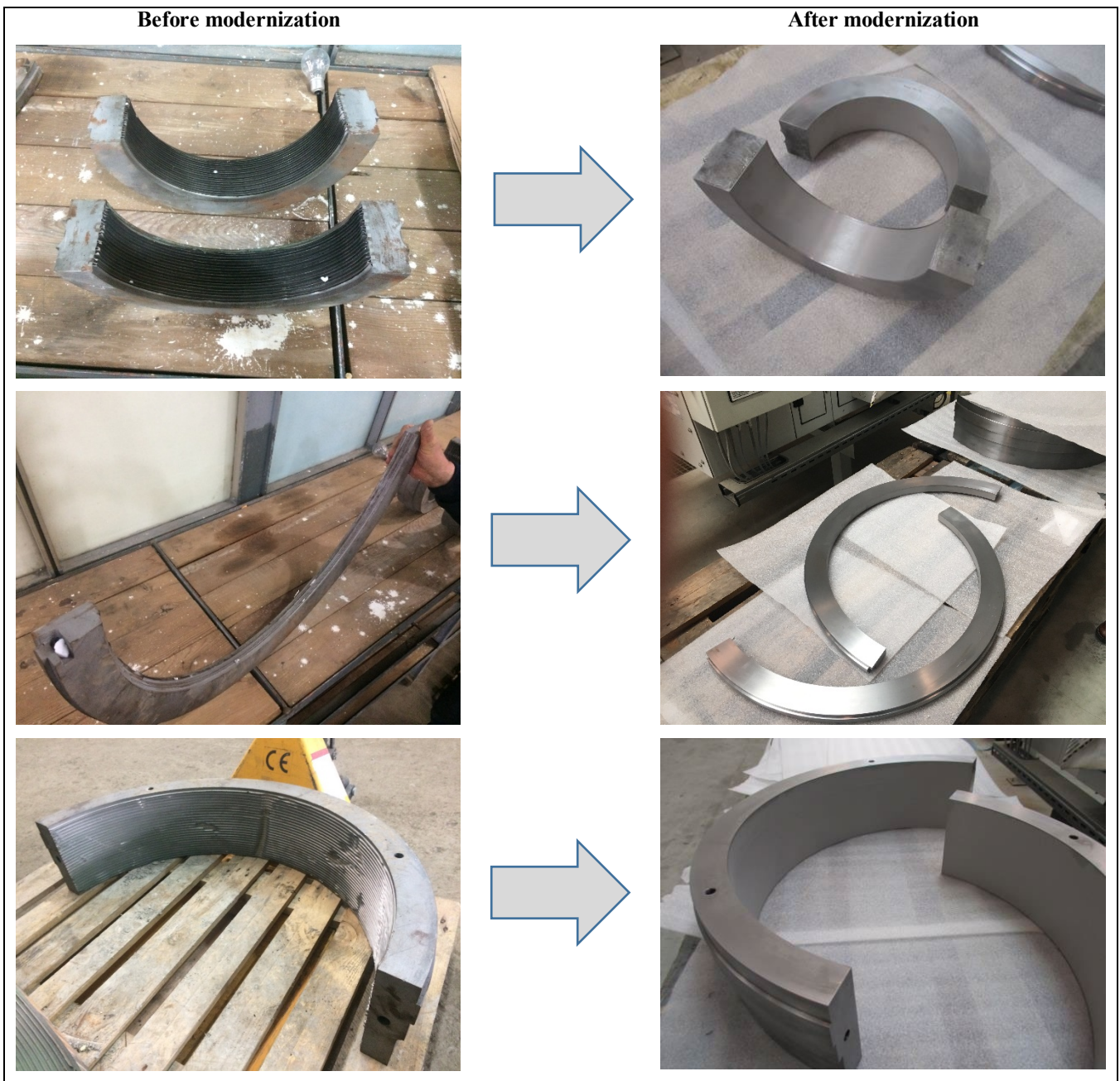


Figure III. View of the sealing elements before and after modernization

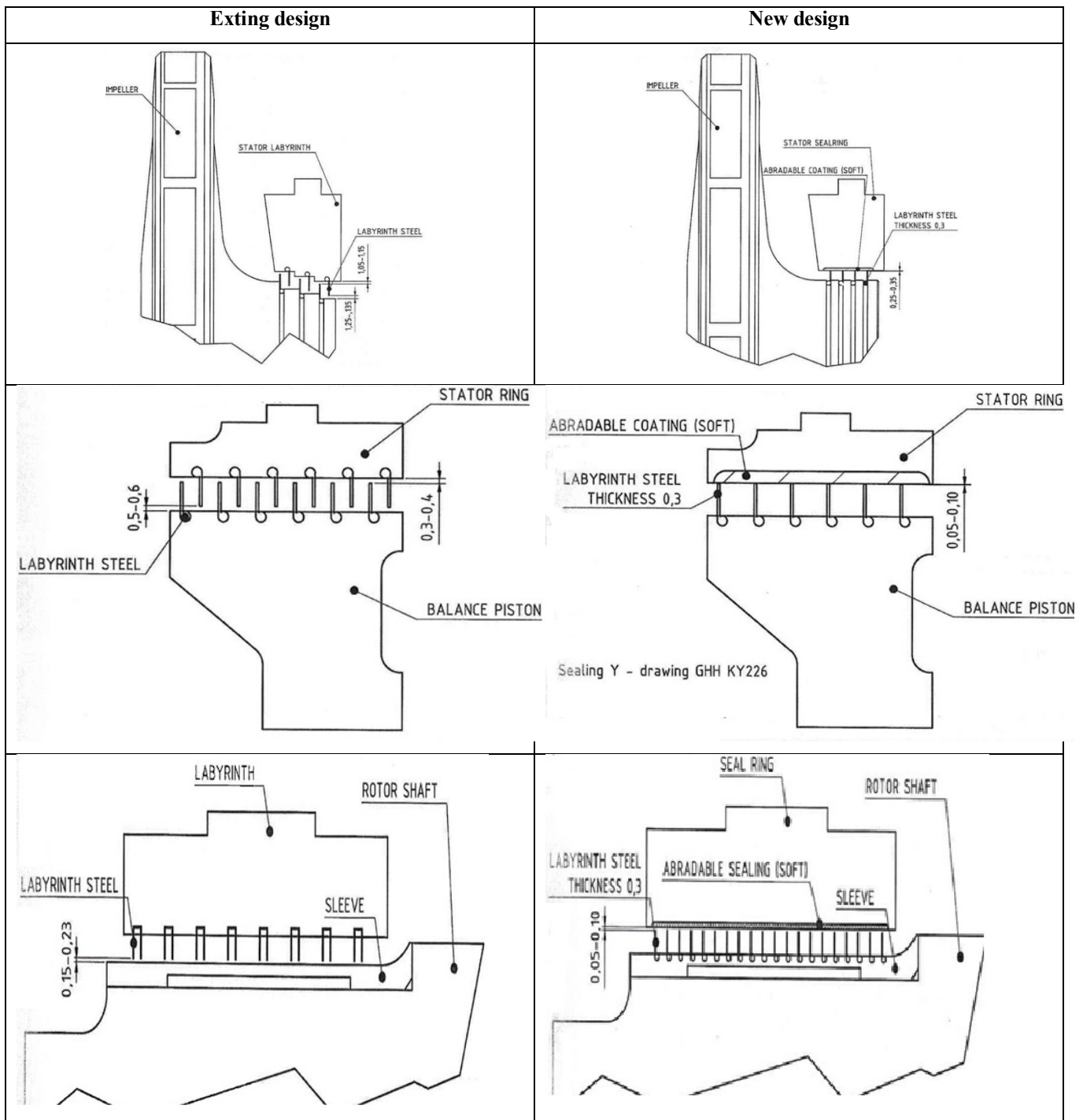


Figure IV. Manufacturing drawings of old and newly designed sealing elements

When compressor air leaks are reduced, compressor pressure-flow increases, when the flow rate increases, ammonia/air percentage composition is preserved in ammonia incinerators, and more airflow is fed for the same ammonia flow rate. Therefore, the ammonia and oxygen reaction conversion rate increases, and more nitrogen monoxide (NO gas) is produced. As more NO gas directly affects the oxidation and absorption reaction in other steps, nitric acid production also increases. Therefore, the compressor outlet pressure is one of the most critical factors affecting the efficiency of the compressor. Turbo compressors formed with impellers arranged on a single shaft, but if the capacity is high, it is efficient.

For example, in screw compressors, although the pressure chamber and the air suction chamber are not separated by definite boundaries, air leaks are still very low. These are high speed compressors. In low-capacity compressors, the yield decrease in the final stages is unacceptably large. High outlet pressure production means extra energy consumption of the compressor, that is, additional operating cost. Pressure levels that are accidentally set higher than necessary will result in compressor control set values (minimum and maximum) that can be re-examined and gradually reduced to the required levels not to damage the sensitive equipment used in the business [9,10].

However, most of the time, the compressor outlet pressure level increases to compensate for the pressure losses between the compressor and critical end-use points in the system. These losses, which create the need to use compressed air like artificial equipment in the compressed air system, cause low system performance and unnecessary energy consumption of the compressor. In compressed air systems with any airflow restricting factor, it is imperative to increase the system pressure to achieve the pre-calculated flow capacity.



If the pressure losses from the compressed air outlet to the end-use point in a compressed air system do not exceed 10 % of the compressor outlet pressure, it can be said that this system is adequately designed. Since the inlet conditions are known during these calculations, it is easy to design the suction edge. However, in design calculations based on the pressure edge dimensionless mass flow parameter, the losses must somehow be included in the calculations. The view of the turbo-compressor before and after rotor maintenance, coating process, and high-speed balancing is shown in Figure V.



Figure V. Before and after rotor maintenance, coating process and balancing adjustments.

As a result of modernization, the compressor pumps air at a higher flow rate than before, and thus, the amount of acid produced increases. Likewise, an increase in airflow and acid production was achieved according to the design values. The test results before the modernization of the compressor are given in Table I, and the test results after the modernization are shown in Table II.

Table I. Compressor test results before modernization

Speed (rpm)	Air (m <sup>3</sup> /h)	Temperature (°C)	Production 100% acid (ton/day)
5250	89375	22	590
5250	87690	21	580
5250	87550	22	585
5250	87185	22	580
5250	87760	22	585
5250	87720	20	585
5250	88685	18	590
5250	87995	21	585

The pressure ratio produced by the impellers is proportional to the square of the operating speed. Therefore, unbandaged impellers can make much higher pressure ratios than bandaged impellers. However, non-bandaged impellers tend to be less effective due to the high losses associated with wingtip leakage flow. In addition, there is no tip leakage in a bandaged impeller. Therefore, while the centrifugal compressor is pre-designed, the head, flow, and speed are taken as a basis.

Table II. Compressor test results after modernization

Speed (rpm)	Air (m <sup>3</sup> /h)	Temperature (°C)	Production 100% acid (ton/day)
5244	97358	18	680
5252	97429	19	670
5246	96830	18	670
5250	97613	17	670
5241	97337	18	674
5247	97010	19	674
5246	96565	20	673
5247	97163	18	673

An unbalanced rotor will vibrate at the frequency of shaft rotation speed due to the unbalanced mass's centrifugal force. Therefore, a machine with an unbalance condition is expected to produce a sinusoidal sine wave and a corresponding dominant peak in the spectrum at the shaft rotation speed. Turbo-compressor balance measurement values are given in Table III.

Table III. Turbo-compressor balance measurement values

Speed (rpm)	Pedestal 1 V		Pedestal 2 V	
	mm/s, rms	mm/s, rms	mm/s, rms	mm/s, rms
2200	0.084	246	0.055	226
2700	0.053	112	0.084	114
4900	0.060	77.9	0.132	54.1
5600	0.278	101	0.151	26.9

Turbo-compressors are continuous service compressors, with the advantage of having very few moving parts, especially in applications where high airflow is required and mainly where oil-free air is required are suitable for use. However, unbalance forces put pressure on bearings and seals, exacerbating looseness problems and can trigger resonances. The force created by an unbalance weight is related to the square of the velocity, so high-speed machines can generate enormous unbalance forces and therefore cannot be allowed to go out of balance. In case of exit, it will cause inevitable damage. The balance report chart of the turbo-compressor is given in Figure VI.

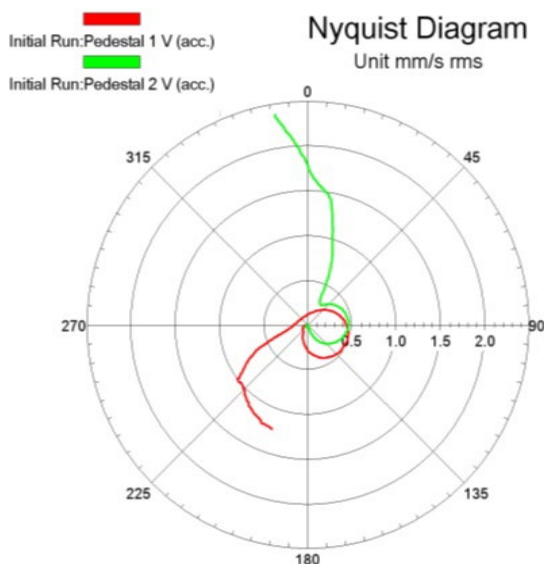


Figure VI. Turbo-compressor balance report chart

#### IV. Conclusion

In the production of dilute nitric acid, it was observed that the amount of air produced by the turbo-compressor, which sends air to the oxidation reactor, determines the acid production capacity. As a result of the modernization, there were increases in air produced by 6.41% and 7.69%, respectively, compared to the previous two years. In addition, there was an increase in acid production by 10.22% and 7.61%, respectively. Compared with the design values, a rise of 2.87% in air and 4.26% in acid production has been achieved after modification. In terms of economic gain by increasing the efficiency of the turbo-compressor, it has been calculated that the cost of modernization pays off in about 90 days.

#### References

- [1] Zhang Chaowei, Dong Xuezhi, Liu Xiyang, Sun Zhigang, Wu Shixun, Gao Qing, Tan Chunqing, "A method to select loss correlations for centrifugal compressor performance prediction", *Aerospace Science and Technology*, Volume 93, 2019.
- [2] Xiao He, Xinqian Zheng, "Flow instability evolution in high pressure ratio centrifugal compressor with vaned diffuser", *Experimental Thermal and Fluid Science*, Volume 98, 2018.
- [3] Fei Chu, Fuli Wang, Xiaogang Wang, Shuning Zhang, "Performance modeling of centrifugal compressor using kernel partial least squares", *Applied Thermal Engineering*, Volume 44, 2012.
- [4] Waleed Albusaidi, Pericles Pilidis, "An Iterative Method to Derive the Equivalent Centrifugal Compressor Performance at Various Operating Conditions: Part II: Modeling of Gas Properties Impact", *Energies*, Volume 8, 2015.
- [5] Hong Zhang, Hang Zhang, Zhuo Wang, "Effect on Vehicle Turbocharger Exhaust Gas Energy Utilization for the Performance of Centrifugal Compressors under Plateau Conditions", *Energies*, Volume 10, 2017.
- [6] Xiang XUE, Tong WANG, Tongtong ZHANG, Bo YANG, "Mechanism of stall and surge in a centrifugal compressor with a variable vaned diffuser", *Chinese Journal of Aeronautics*, Volume 31, Issue 6, 2018.
- [7] Nobumichi Fujisawa, Tetsuya Inui, Yutaka Ohta, "Evolution Process of Diffuser Stall in a Centrifugal Compressor With Vaned Diffuser", *J. Turbomach.* Volume 141(4), 2019.
- [8] Klaus Brun, Sarah Simons, Rainer Kurz, Enrico Munari, Mirko Morini, Michele Pinelli, "Measurement and Prediction of Centrifugal Compressor Axial Forces During Surge—Part I: Surge Force Measurements", *J. Eng. Gas Turbines Power.*, Volume 140(1), 2018.
- [9] Charles Stuart, Stephen Spence, Dietmar Filsinger, Andre Starke, Sung In Kim, "Characterizing the Influence of Impeller Exit Recirculation on Centrifugal Compressor Work Input", *J. Turbomach.* Volume 140, 2018.
- [10] Bachir NAIL, Abdellah KOUZOU, Ahmed HAFIFA, Ahmed CHAIBET, "Parametric Identification and Stabilization Of Turbo-compressor plant based on matrix fraction Description using experimental data", *Journal of Engineering Science and Technology* Volume 13, 2018.

# Prediction of cover bunch quality after harvesting period of Tunisian date palm

Wafa Guedri  
Laboratory of Textile Engineering,  
University of Monastir  
Monastir, Tunisia  
wafa.guedri@gmail.com

Mounir Jaouadi  
Higher Institute of Technology Studies  
ISET (Ksar Hellal)  
Monastir, Tunisia  
jy.mounir@gmail.com

Slah Msahli  
Higher Institute of Technology Studies ISET  
(Ksar Hellal)  
Monastir, Tunisia  
slah.msahli@gmail.com

**Abstract**— Bagging date fruit is a necessary practice in the date palm. It is used in the date crop to protect fruits from humidity, rain and insects. At present, it is focused to many studies to improve its performances. In this work, a new method is proposed intending to solve this multi-criterion phenomenon. A practical mathematical tool named desirability function allowing the prediction of the cover bunch is developed. This approach permitted to define the global quality of the bag through a global quality index. In this work, five bagging products are selected to be inspected, used for the first time to protect date fruit against carob moth and rain. Find results allowed the identification of ideal cover for farmers to satisfy their needs that would has the highest properties.

**Keywords**—Desirability, quality, date fruit, nonwoven, satisfaction

## I. Introduction

Date palm is the most significant fruit crop grown in arid and semi-arid regions of Middle East and North Africa. In Tunisia, date palm is the major factor of oases farming. It represents the main financial resources of farmers since date fruits are used for food or other commercial purposes [1]. For occurrence, about 10% of the Tunisian population is dependent, on date palm and its related crops [2]. However, this important crop is actually in danger by severe factors. A majority of the date cultivars are susceptible to pests mainly carob date or yield poorly caused by autumnal rain, sunburn and wind. These factors had led to reduce the date quality, which menace the physical aspect of the date fruit. Hence, it is very important to elaborate a strategy intending to improve the quality of the Tunisian date fruit. Covering procedure of date bunch is regarded as the most important practice for date palm to get a date with better quality and an economical yield. It offers numerous advantages and is used in the date fruit cultivated in order to protect date from rain, high humidity, bird and insects [3]. Different covers are existing for date fruit protection [4].

The requirements for the cover date bunch were mentioned by D.E bliss in his study [5]. His study conducted that the best cover must be water-resistant during heavy rain and allows maximum aeration because the vapor transpired by the date surface is imprisoned by the bag and guides to water injury and infection. However, there are a few works when evaluating the value of the bag used for protecting date fruit [5]. The first work on the measurements of cover bunch features was described by Denis [6]. He planned the first cover in the shape of cloth bag based on flexible woven

textile permitting protection from wind, insects and heavy rain. In fact, it is important for Tunisian farmers to be sure that cover bunch will keep its waterproofing and remain its strength during maturity period.

Previous works confirmed that using traditional covers like plastic film, mosquito net and kraft paper had a prominent effect on quality and yield of date fruit [5]. The ideal bag requests a compromise between several requirements of Tunisian farmers for efficient date protection. In fact, the satisfaction of Tunisian farmers is a phenomenon that requires satisfaction of a set of features during harvesting period.

In this work, we propose a mathematical approach in order to estimate the Tunisian farmer's satisfaction using an index named global quality index "QI". The desirability functions are used to develop this index. Modeling this satisfaction allows developing the ideal choice of cover date bunch.

## II. Material and methods

### A. Selected samples

During the 2016-2017 seasons, three mature date palm trees, received the same cultural practices at the experimental plot of Institute Arid, "Atilaat" (Jemna, Kébili) were selected for this experience. The number of bunches was adjusted to six per tree. The covers were as follows: bunch cover with nonwovens N1 (polyester high weight), N2 (polyester low weight), N3 (polypropylene high weight), N4 (medium weight) and N5 (low weight), no bagging (control). Bunch covers with different products (Table I) was happening three months before the collect and continued during the full maturity period.

Table I. Properties of the selected nonwoven samples

Properties	Nonwoven samples				
	N1	N2	N3	N4	N5
Supplier	German	German	German	German	German
Composition	Polypropylene	Polypropylene	Polyester	Polyester	Polyester
Color	White	White	White	White	White
G (g/m <sup>2</sup> )	60.3	41.36	75	50.2	34.9
T (mm)	0.44	0.38	0.34	0.31	0.17
Gr-mass per surface unit, T-Thickness of the bags					

Five nonwoven samples were selected from the textiles offered by Freudenberg Group, plastic film and mosquito net from GIfruit-Tunisia. Figure I show the five samples of

covers chosen to study the protection efficiency of date fruits against rain and the carob moth.

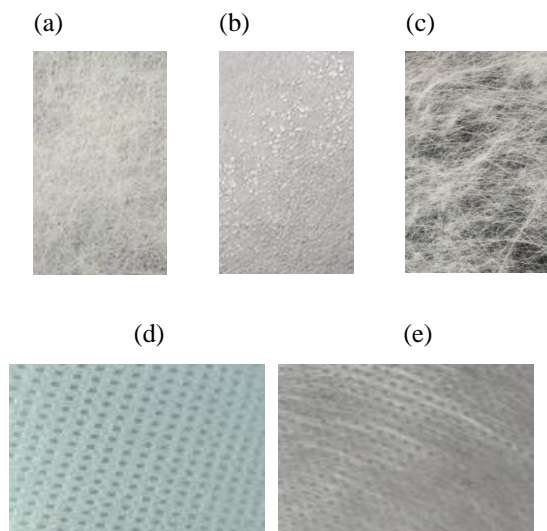


Figure I. Images of used bagging samples: (a)N1; (b)N2; (c)N3; (d)N4; (e)N5;

## B. Development of quality index

Development of the nonwoven covers quality is divided to four steps:

1. Choice of properties
2. Property conversion to a common value
3. Attribution of weighting
4. Aggregation of indices

With a focal point on developing the global quality index QI, we have pursued the procedure offered by figure II.

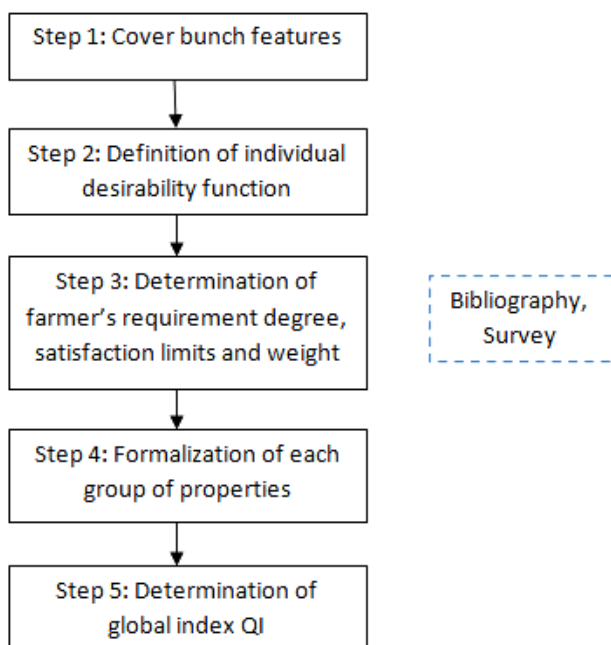


Figure II. Design of quality index QI

### B.1. Choice of parameters

The arrangement of the quality of cover date bunch is hard for all features that define these products. It is then necessary to select a set of parameters to reflect together the global quality of the bag for a particular use [9]. The selection of the properties is accompanied by a uncertainty and subjectivity which are necessary for the utility of all index. The selection of properties is based on data collected from Tunisian farmers, organizations and researchers troubled with the protection of dates.

In this work, the quality of the cover date bunch is characterized by multiple parameters expressing the ability of these bags to be satisfied. There are many physical and mechanical properties that influence this quality. The number of required parameters is fixed for all bagging products. There is a set of parameters that the covers must have to obtain it suitable for the protection of the date bunch. The cover date bunch must be flexible and maintain the movement of the date regime, so that tearing is the most important parameters for any type of covering product.

Since a cover date bunch is mainly used to keep out insects, the bag must have features of suitable dimensions to surround the entire bunch. Textile strength is a key property related to the features of fabric durability. Since the cover bunch is focused to multidirectional forces, the tensile and tear strength are used to determine the strength of the product. The properties of permeability and resistance to water penetration are quality parameters that establish the effectiveness of the bag and the preservation of the quality of date fruit.

### B.2. Parameter conversion to a common value

The properties of the cover date bunch are expressed by different units. In fact, levels can differ from one property to another. The classification of cover quality intended to be formulated. All properties must be changed to a common value that is a quality index "QI". The mathematical equations that convert the values of the properties into indices are made according to the desirability function. An index will be in a scale of 0 to 1 [10]. If cover quality properties get together the set specification values, a value of 1 is selected [11].

In the equations' development, linear functions are made as a correlation between cover properties and the quality of the cover. In order to convert cover quality properties to a common value, the lower and the upper limits of each property must be identified. The upper and the lower limit useful for the bag quality properties is based on the theoretical knowledge, the specifications and the experts opinions' who participated in the survey.

In this approach, the quality properties of the cover date bunch are divided into three mathematical functions named Derringer and Such desirability.

The first functions are monotonous and can formalize a preference in terms of decrease or increase. Harrington [11] calls them One-sided. One-sided desirability characteristic has three parameters. We get the example of an increasing function. They are each parameterized by two points named lower limit  $Y_1$  (upper limit  $Y_u$  for decreasing function) and target value  $Y_T$ . The first  $Y_1$  indicates the lower region at which the values of  $y_j$  are considered unacceptable; the second  $Y_T$  designated the area of satisfaction.

Finally, the parameter  $s$  makes it possible to define the curvature of the function. In fact, the linear function corresponds to  $s = 1$ . Depending on the  $s$  values, the curvature can be oriented.

We employ the desirability function to maximize a property, where “ $d_j$ ” is calculated according to equation 1 [12].

$$d_j = \begin{cases} 0 & \text{if } Y_L \leq y_j \\ \left( \frac{y_j - Y_L}{Y_T - Y_L} \right)^{s_j} & \text{if } Y_L \leq y_j \leq Y_T \\ 1 & \text{if } y_j \geq Y_T \end{cases} \quad (1)$$

We use equation 2 to minimize a property.

$$d_j = \begin{cases} 1 & \text{if } y_j \leq Y_T \\ \left( \frac{y_j - Y_U}{Y_T - Y_U} \right)^{s_j} & \text{if } Y_T \leq y_j \leq Y_U \\ 0 & \text{if } y_j \geq Y_U \end{cases} \quad (2)$$

The last function proposed by Harrington [12] is called Two-sided and permitted the targeting value for  $y_j$ , or specifying a satisfaction region. It is situated by two points, the lower  $Y_L$  and the upper  $Y_U$ . In addition,  $Y_T$  value is used to put the incline of the function, where “ $d_j$ ” is calculated according to equation 3.

$$d_j = \begin{cases} 0 & \text{if } y_j \leq Y_L \\ \left( \frac{y_j - Y_L}{Y_T - Y_L} \right)^{s_j} & \text{if } Y_L \leq y_j \leq Y_T \\ 1 & \text{if } y_j \geq Y_U \\ \left( \frac{y_j - Y_U}{Y_T - Y_U} \right)^{s_j} & \text{if } Y_T \leq y_j \leq Y_U \end{cases} \quad (3)$$

### B.3. Attribution of weightings

After the checking of the properties, weightings are assigned to each property specified its importance in the global quality of the cover date bunch. The range of values for these weights varies from zero to ten. The mean response is employed to compare the importance of each property. The value is calculated according to the Equation 4 [13]:

$$M = \frac{\sum O \times V}{N} \quad (4)$$

Where M: Mean response; O: occurrence; V: market value and N: number of response

The requirement degree of the Tunisian farmers is noted “ $s_j$ ” and calculated using the equation 5. We consider that if  $s_j = 1$ , the requirement degree is average. On the other part, if

$s_j > 1$ , we believe that the requirement of the farmer is too high, and if  $s_j < 1$ , we can consider that the farmer’s requirement is too low [15].

$$s_j = \frac{RV}{SL} \quad (5)$$

Where RV: average recorded value; SL: average scale value

In Equation 6,  $w_j$  is the relative importance related to the response  $j$ . The relative importance  $w_j$  is a comparative level for weighting of the individual desirability  $d_j$  in the global desirability cover and it varies from the least important ( $w_j=1$ ) to the most important ( $w_j= 5$ ) [14].

$$w_j = \frac{RV}{10} \quad (6)$$

### B.4. Aggregation of indices

The presence of different criteria and the absence of a relationship solutions guide to the use of strategies that consider these particularities. The method we study in this work consists in transforming the multi-criteria difficulty into a single criterion problem. Geometric mean aggregation in Equation 7 and arithmetic mean aggregation in Equation 8 are used to join individual desirabilities “ $d_i$ ” in overall desirability  $D_g$  calculated by each group of property. The weight of each property is primordial in the determination of the  $D_g$ .

Finally, different  $D_g$ 's are collected together in a degree of global satisfaction defined by QI using the desirability function of Derringen and Suich [16].

$$D_g = (d_1^{w_1} + d_2^{w_2} + d_3^{w_3} + \dots + d_n^{w_n})^{\frac{1}{\sum w_j}} \quad (7)$$

$$D_g = (d_1^{w_1} \times d_2^{w_2} \times d_3^{w_3} \times \dots \times d_n^{w_n})^{\frac{1}{\sum w_j}} \quad (8)$$

The quality index values QI vary between 0 and 1. The 0 value designated any satisfaction and the value of 1 corresponds to an entire satisfaction for the overall quality of the cover bunch.

## III. Results and discussions

When bag is detached from the date palm, some properties must be calculated. A full description of the analytical methods employed is ready to study the status of covers removed.

### A. Cover bunch features after bagging period

Experimentation of the different covers used is carried out at the TTS laboratory in Moknines. The requirements of the covers must be related to the application on the date palm. Experimental results presented in the form of pictures indicate a lack of tear resistance of N1 and N2 samples. The literature indicated that the problem was due to the low UV resistance of polypropylene. Figure II shows that the covering period causes a significant destruction of the N1 material while there are partial tears of the N2 material. In contrast to the N1 and N2 materials, the N4 material has retained its strength for 3 months.





Figure III. Tears of the N1 material after the bagging period

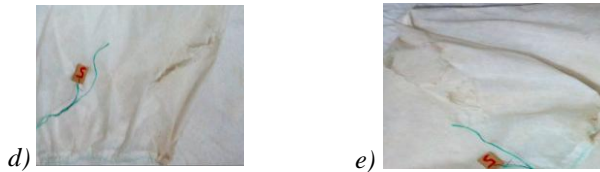


Figure IV. Tears of the N2 material after the bagging period  
Global desirability function of Derriguer and Suich is used to calculate the QI were values are shown in Table II.

Table II. Quality index values

Samples	N1	N2	N3	N4	N5
Dg (structure)	0,42	0,12	0,25	0,56	0,43
Dg (resistance)	0,1	0	0,11	0,32	0,18
Dg (protection)	0,18	0,27	0,05	0,3	0,09
QI	0,007	0	0,001	0,05	0,007

Evaluation of covering product qualities by Tunisian farmers has been performed in order to determine farmer's requirement degree and weight for each cover bunch property. Measured properties have been transformed to individual satisfaction degree and global desirability index was determined for the appreciation of farmer's satisfaction. The quality index of all samples were determined to confirm the ability of these materials to protect against violence wind, rain and to limit dust obtained for N1 and N2 materials. In the case of N1, N2 and N3, their bad resistance after the period of bagging deeply depreciated their overall quality.

#### IV. Conclusion

In this work, we have examined the multi-criterion phenomena of cover date bunch, during maturity period includes different properties simultaneously. A desirability method for appropriate nonwoven covers selection based on Derriguer and Suich function was presented. This approach is suitable to minimize, maximize and achieve target values of some objective functions at the same time. It is used to convert the different output in one output that is the set of the individual output while affecting different weights according to the importance of each property in the studied case. This method has been applied to reduce the different requirements affecting the quality of the cover bunch in one index representing the global quality of the nonwoven bags varying between 0 and 1. This process advanced the quality properties in term of the ideal output responses. The considered quality indices have revealed that an enhancement in permeability performance guarantees an increase in product satisfaction.

#### References

[1] Jorg Kruse et al., Optimization of photosynthesis and stomatal conductance in the date palm *Phoenix dactylifera* during acclimation to heat and drought, *New Phytologist*, 2019.  
[2] Ilyasse Izanar et al., Corrigendum to "Hard carbons issued

from date palm as efficient anode materials for sodium-ion batteries", *Carbon* 146, pp 844, 2018.  
[3] S. Sakin, "Analytical methods applied to the chemical characterization and classification of palm dates (*Phoenix dactylifera* L.) from Elche's Palm Grove," University of Alicante, Alicante, 2013.  
[4] Salah Amroune et al., "Investigation of the date palm fiber for composites reinforcement: thermo-physical and mechanical properties of the fiber", *Journal of natural fibers*, 2019.  
[5] Wafa Guedri, Mounir Jaouadi and Slah Msahli, 'Evaluating farmer's satisfaction of different agrotexile cover bunch using desirability function', 2021.  
[6] SP.Denis, "Means and method for protecting Deglet Noor Dates," *Brevet US2001051240*, 2001.  
[7] M. Selmane, Thèse en écologie animale, Annaba, FA faculté de sciences, département de biologie, 2015.  
[8] D. Zohary, Domestication of Plants in the Old World: The origin and spread of domesticated plants in Southwest Asia, Europe, and the Mediterranean Basin, Oxford, 2012.  
[9] J. Smartt, Evolution of crop plants, 2nd edition, 1995.  
[10] P. Munier, Le palmier dattier .coll. Techniques agricoles et productions tropicales, France: Ed. g. Maisonneuve et larose, XXIV, p. 221, 1973.  
[11] A.Hadj taieb, S.Msahli and F.Sakli "Optimization of the Knitted Fabric Quality by using Multicriteria Phenomenon tools," *International Journal of Fiber and Textile Research*, vol. 3, no. 4, pp. 66-77, 2013.  
[12] A. M C Balasooriya, "Development of a comprehensive fabric quality grading system for selected end uses," *National Engineering Conference, 19TH ERU SYMPOSIUM*, pp. 33-37, 2013.  
[13] A. M. Shrahan Kumar Gupta, "Optimization of durability of Persian hand-knotted wool carpets by using desirability functions," *Textile Research Journal*, pp. 1-10, 2016.  
[14] F.Dabbebi and S. B. Abdessalem, "New approach for appreciating the surgeon's satisfaction of braided sutures," *Journal Of Industrial Textiles*, pp. 1-23, 2015.  
[15] Zehdi S, "Analysis of Tunisian date palm germplasm using simple sequence repeat primers," *African Journal of Biotechnology* Vol. 3 (4), vol. 3, no. 4, pp. 215-219, 2004.  
[16] S. Rhouma, "Genetic diversity in ecotypes of Tunisian date palm (*Phoenix dactylifera* L.) assessed by AFLP markers," *The Journal of Horticultural Science and Biotechnology*, vol. 82, no. 6, p. 929-933, 2007.  
[17] D. Mostafa, "Effect of Bunch Bagging on Yield and Fruit Quality of Seewy Date Palm under New Valley Conditions (Egypt)," *Middle East Journal of Agriculture Research*, vol. 3, no. 3, pp. 517-521, 2014.  
[18] W.Guedri, M.Jaouadi and S.Msahli, "New approach for modeling the quality of the bagging date using desirability functions," *Textile Research Journal*, vol. 86, no. 19, p. 2106-2116, 2016.

# Analysis of the Waterfront Transformation of the 'Plazh' Area of the City of Durrës, Albania.

Anna Yunitsyna  
Department of Architecture  
Epoka University  
Tirana, Albania  
ayunitsyna@epoka.edu.al

Mirela Hasanbashaj  
Department of Architecture  
Epoka University  
Tirana, Albania  
mhasanbashaj10@epoka.edu.al

**Abstract**— In many waterside cities coastline urbanization is an important issue for architects and urban planners. Albania has been unsuccessful in developing a balance between the demand that the public has to access the coastline and the land use. The coastline nature, has changed because of demographic, hasty urban and economic growth. This process took place without taking into consideration the social and cultural values of the coastline. Massive constructions have damaged the sewage system causing continuous over flooding and reduced the coastline greenery. This study investigates the transformation of the seaside of so-called 'Plazh' area during the last 20 years. The study is divided into the 5 years periods, in order to have a more complete overview of the urban changes. The types of buildings included in the study are residential and commercial. In this research, have been used the information collected from the site surveys and observations and archival. T this study presents a report explaining the site problems and transformation of the costal line of Durrës during years, a comparison between the master plans, proposed regulatory plans and the evaluation of the actual situation.

**Keywords**— *coastline, urbanization, land use, height limit, urban regeneration, waterfront*

## I. Introduction

City of Durrës is a coastal town which lies in the western part of the region. It holds the largest port in Albania which makes it an important landmark for tourism. It is the second largest city from the politics, economy, administration, education and culture standpoint. Durrës itself holds good opportunities, lying in a strategically position between north and south of Albania. It is the trade center of intersections and has won important functions and intensive territory usage, becoming one of the most developed regions of the country. The purpose of this work is to track the transformation the coastline of the city of Durrës and to highilht the urban problems emerged during the last 20 years. The demand, inspiration and decision of the investors and owners for their private property have played an important role on the transformation of city and its coastline in terms of land use and urban planning. The study is focused on the analysis of so called 'Plazh' area using the building location, building height and land use criteria. The study examines the buildings of post socialist period, specifically through 1990-2015. The coastal line has a combination of elements such as residential housing, high rise buildings, clubs, restaurants, shops etc. Redeveloping the water front means re-utilizing the urban lands that were left behind through the years.

The major problem comes from the reclaimed land which changes the relationship between the city and the waterfront. In most of the cases the lands that were reclaimed have been restricted to the general public by the private users. Durrës has evolved with the land reclamation but such development has a negative result with the urban waterfront development with the rest of the region and Europe [1]. Land reclamation in

Albania has led to a planning policy in order to provide economical and suitable land for building. The government is seeking to accommodate public projects whereas the private ownership look for a land for development. In the past years, people in Durrës have expressed their frustration that they have with the inaccessibility to the sea. The political and economic changes that occurred after the 90s, were followed by large flows of migration from rural areas, to urban areas. The urban development of the territory, were largely the product of free initiative and spontaneous development reflected in the creation of the first informal neighborhoods. All the economic activity was concentrated in the central and coastal areas of Albania, but local authorities were unable to provide appropriate urban regulatory plans and to control the process of construction [2].

## II. Waterfront Development Principles

In ancient times, societies used to live in waterfront areas, such as next to Tigris, Nile and Euphrates. In order to sustain life and satisfy the biological dependency on water, humans historically needed to locate near fresh water [3]. Waterfront means 'the part of a town or city adjoining a river, lake, harbor, etc.' according to the Oxford American Dictionary of Current English [4]. 'Waterfront is the urban area in direct contact with water' [5]. According to Moretti, port activities and infrastructures occupy waterfront areas. Waterfront is defined as an interaction area between water and urban development [6]. Although in the vocabulary the meaning of waterfront is clear, in the literature, it is met by using different words instead of the term waterfront such as: riverside, river edge, water edge, riverfront, a city port, and harbor front [4], [7]. Waterfront identifies the water's edge in urban areas [8]. The water body may be 'a creek or canal, river, lake, ocean, bay,' or even artificial [9]. To sum up, the waterfront area is a confluence area of water and land. It is the edge of land and also the edge of water. The waterfront is found as a continuous process in most places where settlement and water are juxtaposed, whether or not a commercial port activity is or was present. It is an area that has a high density of elements and activities that affect each other. In geographical aspect, the urban landscape is a synthesis of climate, soil, biology and physiognomy. An integrated landscape is a combination of natural landscape and artificial landscapes including architecture, streets and squares [10].

In most of European cities, the development process of urban waterfront area is like the following: prosperity, decline and re-development. Urban waterfront areas flourishing timewasbefore 1920s. Before the industrial revolution, there was a development of the society, there was a people dependence on natural water sources, and the water was used for every day live but also for travel. With the development of water traffic, waterfront areasbecame very important for their cities because trading development began and it had a positive impact in the citizens economy, so waterfront areas became

the centers for many people's lives. However, there was a decline of urban waterfront areas from the 1930s to 1960s. After the industrial revolution, there was a rapid growth of the population; along the water were located many industrial and transport companies in order to get more benefits. This brought water pollution because much sewage rushed into water so inhabitants did not want to live there anymore [11]. A reconfiguration of the water aesthetics is done by urban waterfront regeneration. The way in which the waterfront looks actually depends on the creation of attractive waterside open spaces, besides the provision support of residential and business space with waterside prospects.

Waterfront cities expand over reclaimed land, and there is a limited number of studies relating to the waterfront buildings expansion. The waterfront area is where the land interacts directly with the sea. It is a dynamic place which is totally open to the action of waves, wind, currents and tides that can expand the shore with sedimentary deposits and also erode it. It is very important to provide the vision of the future sustainable development of the waterside urban edges and to improve the social, aesthetical, physical and economic conditions of those areas [12]. Besides of the engineering tasks the expansion of the waterfront should involve systematic and careful planning, the strategy of development and sustainable management and also should consider the interests of people living in the area [13]. The common strategy is to provide the guidelines for the waterfront development which include detailed land use principles and architectural design approaches, such as building heights and the built forms, presence of the public areas, access to the water, human-centered design of spaces, lighting and signage [14]. Waterfronts are usually in the center of conflicts between different actors and parties, they combine the complex environmental problems such as pollution and flood, social issues which is a contradiction of local inhabitants which can belong to the poor groups and the tourists, the intention of the developers to densify the constructions and demand for the public spaces [15].

### III. History of the Development of 'Plazh' Area of Durres

The city of Durrës is positioned in the western part of the county. It has a geographical location with a latitude of 41° 19' on the north, and longitude 19° 27' on the east. It has a total area of 432 km<sup>2</sup>. It is bounded the district of Tirana which is 35 km away, and in the south with the district of Kavaja. As mentioned above, the western boundary of the city is defined by the Adriatic Sea Coastline, and is 30 km long. Durrës is mostly composed of large landfills and hills, which stand 89 m above the sea level, whereas in the city, the average height is only 2 m above the sea level. The establishment of Durrës began on the 7 century B.C. The modern city is built on the ruins of Epidamn which is known as the old city and has there are actually a number of significant elements of the historical and cultural heritage. Durrës Bay is located between Selita cape and the cape of south Durrës. On the entrance part of the capes, the bay shores are higher, while on the inside a smooth slope is seen. On the north part of the shoreline the port of Durrës is located, while on the north east and south east of the port lies the city of Durrës. The suburb of the city has changed a lot through the centuries, which means that also the coastline has undergone many transformations due to earthquakes, which

have occurred in this area [16]. The oldest part of the city, is situated on the hill, overlooking the Adriatic Sea in the west, the bay of Durrës in the south and the plain in the east and north-east. The city is growing in all of the directions, but mostly on the coastline. During the last 20 years the city has suffered a drastic transformation. Nowadays on the 4.5 km of coastal line there are problems with the buildings heights, density, distances and land use.



Figure I. The Regulatory Plan of 1942, Detailed Plan for the 'Plazh' area

In the early 20s the 'Plazh' district was a farmland [17]. In the mid-1930s, a regulatory plan was requested for the city of Durrës, the implementation of which turned Durrës beach before the Second World War in a touristic and recreational area. The royal court, in the mid 1930's charged Durrës municipality and a group of technicians to draft the regulatory plan for the coast. A 4 km length of the coast line was divided in 300 parcels with a 400-500 m<sup>2</sup> each. These parcels were sold to different individuals provided that the building would not be more than 2 story high, 80 meters away from the sea line and half of its surface had to be used as greenery. The first regulatory plan of Durrës is done in 1942 by the architect Leone Carmignani (Figure I). The study presented a functional zoning scheme, an administrative map defining the boundaries and the territory junction.



Figure II. Durres Regulatory Plan of 1957

After the establishment of the communist state in 1946, all of the beach villas were nationalized. The buildings were used as a holiday home for workers, and later, some of them converted into guest villas for foreign party leaders or holiday homes for local directors and officers. The same thing happened to buildings that were close to the 'Bllok', in Illyria, while others were reconstructed as small apartments and were rented for two-week vacation from workers families, mainly from Tirana. From 1950 to 1960, based on the principle of industrial decentralization, the first comprehensive regional and urban strategies were developed and applied [18]. The General Regulatory Plan of Durrës was prepared in 1957 (Figure II). During this period the coastline was considered an important natural element which belonged to the public and was still treated with care.

The second Regulatory Plan (Figure III) was done in 1987, covering a territory of 2800 hectares, which in 1999 extended in 7800 hectares by Durrës Municipality. According to this plan, there was not allowed any construction on 'Plazh' area. It was prohibited to use this area for residential or other purposes.

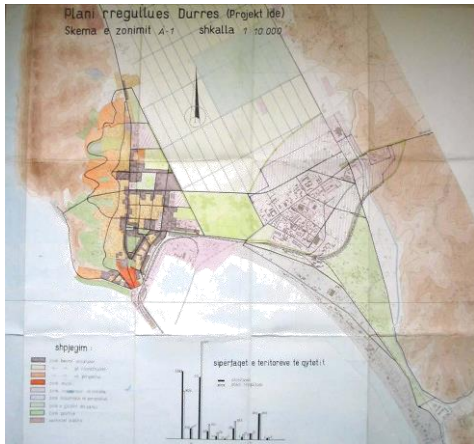


Figure III. Durrës Regulatory Plan of 1987

After the 90's Durrës became one of the most attractive cities with a massive displacement of population from rural areas or little town. This resulted in an urbanization process with exceptional growth rates. The most common phenomena that appeared in Durrës, as also in whole Albania in that period is the process of illegal construction. Another Master Plan was done in 1995 by DAU (Urban and Architecture department, Pescara Faculty). It had orientation purposes, but this Master Plan did not have juridical approval from Durrës Administration. Master Plan for the city of Durrës of 1995 [16], was performed in very beginning of the transformation of the territory and the housing surface that changed the whole structure of the region. The city of Durrës was foreseen to consist of four 'cities' or homogeneous piece of territory: The old historical city; the touristic city; the hills city; Big Natural Park, consisting of hilly coastal park. This Master Plan underestimate the importance that would have the fifth city, the informal one, which under the influx of migrants and immigrants would fill the 'Kënetë' area of the Durrës city.



Figure IV. Satellite image of Durrës, 1995

The Revision of the Urban Conditions of 2005 was prepared by Durrës Municipality. This plan has done the resettlement of the urban criteria of the Coastal area Master Plan (Figure V). It proposed a 'zoning', which was defining only one urban criteria, the height (of 6, 8, and 10 floors).

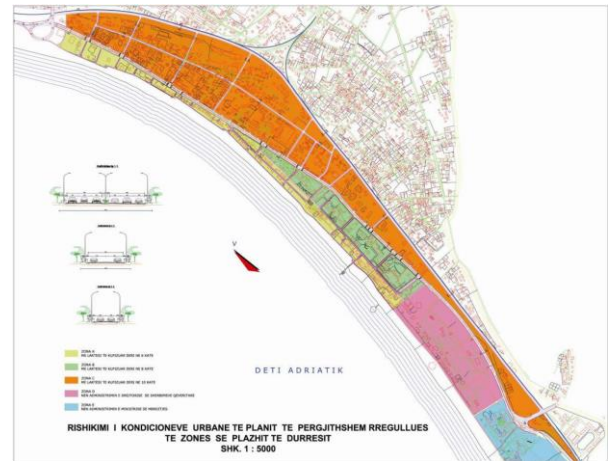


Figure V. Revision of Urban Conditions of the Coastal Area Durrës, 2005

The next plan is designed within the World Bank project on 'Land Management and Administration' during 2007-2011 (Figure VI). The plan extends to 15 years perspective. Durrës is divided in the industrial area, port area, in the city center as well as green areas. The development of industrial area, which includes Porto Romano and Spitallë, is determined by the government. The port is the government property, but it was planned the opening for its residents in order not to remain a closed area as a military base. The city center is conceived as a service area. There is also determined the development of the touristic zone, which will extend the existing part, also in the north of 'Currila' area. The masterplan defined the development of facilities serving the community, such as hospitals, social centers, schools, sports centers and others. Since the territories are already occupied in Durrës, the plan envisages the land owners to secure them. According to this regulatory plan there for area from 'Ura Dajlanit' to 'Plepa' there are 7 land categories which include the criteria of the land use, buildings heights, density of the

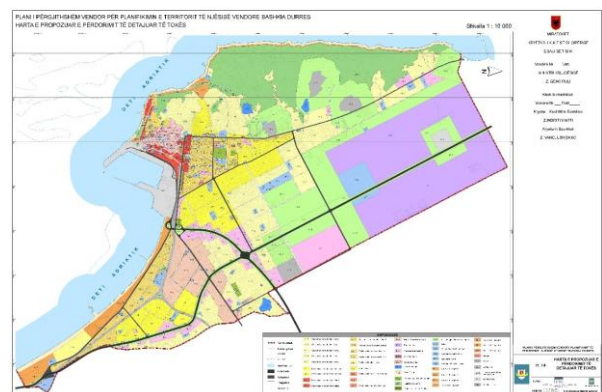


Figure VI. World Bank project on 'Land Management and Administration' for Durrës City, 2011

area and so on.

During the last years the city is experiencing densification, which is characterized by new housing construction as well as high rise and interventions on the existing buildings. The effects of urban densification, limited investments and maintenance underfunding, have accelerated the decline on quality in infrastructure. The new buildings were developed mainly in existing areas with low density

houses, negotiating and compensating owners with apartments and shops.

#### IV. Analysis of the Coastline Development

This comparative and descriptive analysis of coastline transformation from 'Ura Dajlanit' to 'Plepa' is based on study of visual materials such as maps and images, archival research, site survey and site observation.

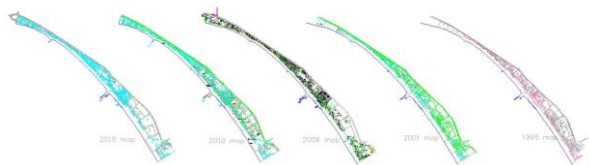


Figure VII. 1995, 2001, 2006, 2010, 2015 maps of Durrës coastline

The basic maps of 1995 (technical drawings, architect Gëzim Hasko, Archive of Durrës, Water Utility) that has been redrawn as a Dwg file, 2001 and 2010 maps (Land Register office), 2006 map (ALUIZNI) and the 2015 map (self-updated on the terrain by using the 2010 and Geoportal Asig as a reference) were provided (Figure VII), and used to analyze the changes of the costal line for every 5 year period. Each of the maps included the information regarding the building heights. Comparison between the maps allows to identify the number of new construction added during each period.

#### A. Number of Construction

After 1990 the number of construction of the buildings has risen year by year. This has completely changed the image of the city, its structure and socio-economic aspects.

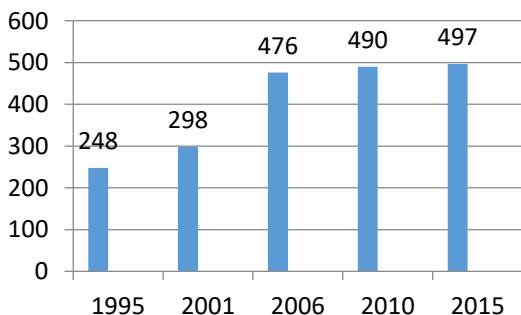


Figure VIII. Total number of buildings per each period

From the constructions analysis graph (Figure VIII), it is evident that in 1995 the total number of buildings was 248. In 2001 the buildings number has reached up to 298 buildings. The highest number of constructions and collaterals is seen from 2001 to 2006 where the number of constructions has reached a total of 476 buildings. From 2006 up to 2010 the rhythm of construction was lower. At the end of 2010 there were 490 buildings in total. On the last 5 years, the number of constructions has decreased. One of the reasons is that the new Regulatory Plan does not allow new constructions on this area, since there is a height building density and there has not remained too much free land to continue with the chaotic construction that has characterized this area. So, at the end of 2015 there are in total 497 buildings. The total number of buildings added is 249. From 1995 to 2001 there were 50 (20

% of the total) new constructions, most of the buildings were 2-3 floor high and mostly small apartments or villas, were used for tourism. The biggest number of new constructions appeared during 2001-2006, when there were constructed 178 (71% of the total number) buildings from which 94 of them are built on the foundations of the existing ones. After 2006, the rhythm of constructions has decreased. There are added 14 (6% of the total number) buildings during this period. From 2010 to 2015 there are constructed 7 new buildings (3% of the total number).

#### B. Building Heights

Durrës was perceived as a low rise city before 1995. However, from this year a large number of large scale developments was constructed along the coastline. First wave of construction was mostly up to five or six floors. The boom development of construction of apartment blocks began from 2001 until 2006. There is a big contrast between the low rise villas with the newly constructed residential buildings. On the maps of the building heights (Figure IX), it is seen that the building heights have increased from 1995 to 2015, but the period which makes contrast is from 2001-2006.

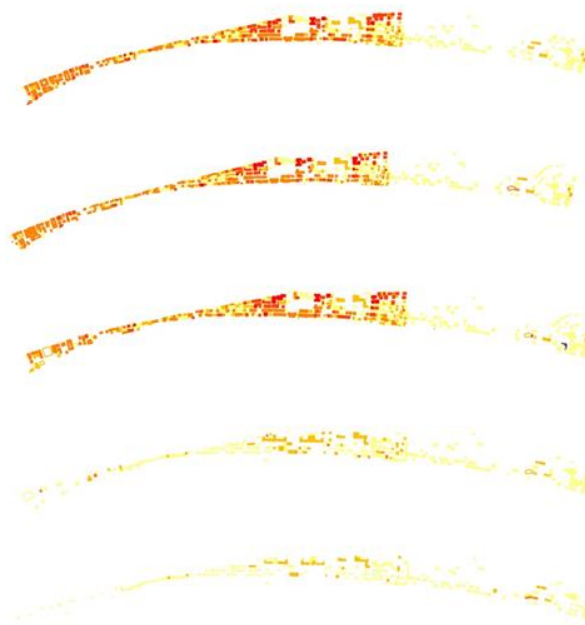


Figure IX. Building heights map of Durrës 'Plazh' area from 1995 (below) to 2015 (above)

The map from 1995 shows that most of the buildings are one up to two story high, so 'Plazh' area had only villas with a max of 3 story high. Those villas were constructed before years 60' and until 1995 there were not constructed any new building because there were not allowed constructions on the touristic areas. On the map of 2001 it is seen that there are not very big changes in comparison with the 1995 map. The new buildings were not high rise and distances between in building were respected. From 2001 until 2006 a massive and uncontrolled expansion of apartment blocks is noted. This period of construction corresponds to the revision of the master plan made in 2005 by Durrës municipality. The revision of urban conditions it is not based in any study done for this area because the last studied master plan was the one made in 1987. Despite on this, the construction continued by

causing many problems. One of the main reasons of the massive construction was land reclamation. In 2005, the Revision of Urban Condition divided the area in 2 parts, the area with buildings heights up to six stores and the area with the building heights up to eight stores. The new buildings that were added were from two up to eight stores. After 2006, the construction continued. Also in this period the buildings were mostly apartment blocks with the heights up to nine floors. From 2010 to 2015 there were not added many buildings. This is because there was not too much of vacant area remained.

The overall height of new buildings increased from 1995 to 2015. In 1995 there were no buildings with 7-8 or 9+ floors. In 2001 there are 7 buildings of 5-6 story height. Furthermore, in 2006 and further on, the number of high rise buildings has drastically increased. Between 2001 and 2006 there are added 24% buildings with 7-8 story high and 7% buildings of 9+ story height. During 2010-2015 just 7 new buildings were added, which refers to the new master plan and the Urban Laws which came in power.

### C. Building Distances and Parcel Use

According to the map of the buildings problems (Figure X), it is seen that there are a lot of problems with the buildings heights, distances from each other and use of the parcels. Red buildings are out of the construction criteria according to the 2005 Revision of the Urban Conditions and the 2011 Regulatory Plan. The pink color represents the buildings added during years that have not respected the distance rule. Most of the buildings should have a blind facade, but according to the site investigation, there are found windows and balconies. There are also problems with the coefficient of use of the parcels. The cyan color indicates buildings with more than 1 problem.

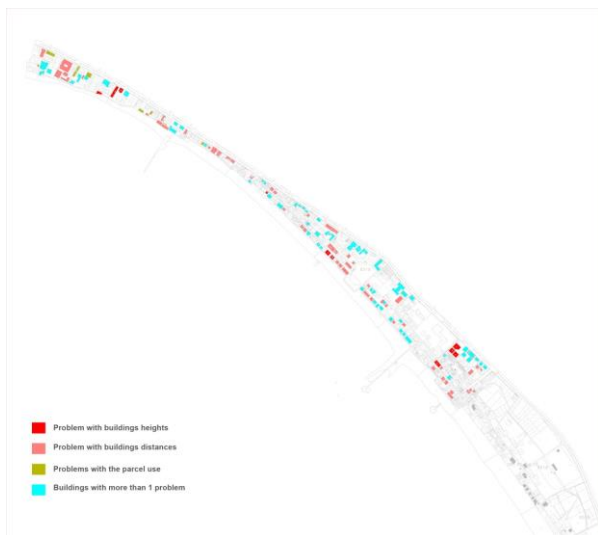


Figure X. Problems with buildings heights, distances and parcels, 'Plazh' area Durres

114 buildings have problems with the distances between them. Approximately 40 buildings have problems with their heights, they exceed the limit of the allowed height, referring to the Revision of Urban Conditions of 2005. 31 buildings have problems with the parcels according to the analyzed map, but since the parcels may have changed during years, the number may be even higher. There are buildings with more than one problem on the map, some of them have distance plus parcels some others may have height plus parcel or distance

plus height. The total number of buildings with multiple problems is 64.

### D. Road Network

The road network of coastline of Durrës has been studied using geographical information systems (GIS). During the last years, the main road which connects the Plazh area with Durrës and Golem got a large increase in traffic and transport demand. The solution to resolve the capacity problem was to provide additional road space, which was the first applied strategy. It was not enough because of the uncontrolled building expansion that occurred these years in the beach area. Secondary roads were not developed. Regarding the existing road system a lack of access in the suburban area is noticed. The main street outlines the area. On the internal area, narrow roads are constructed with dead ends. In the past years, with the massive construction around this area, these secondary roads have been transformed into pedestrian roads. Parking is another major problem since the new buildings did not provide any parking places for the residents. The situation is becoming a major concern to the public because it is getting worse day by day. It also causes air pollution which is one of the biggest problems on the area which damages the tourism and has a negative effect on the quality of life.

The main road is the one which didn't change during the last years. Because of the massive construction a lot of secondary roads were built within the area. They were built

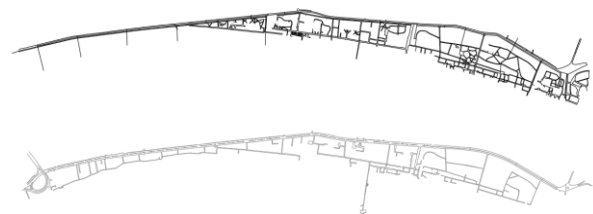


Figure XI. Road network map at 2005 (above) and at 2015 (below)

without a plan which means that there was no infrastructure. The roads were mainly with dead ends. While comparing the two maps it can be seen that before 2005 the secondary roads were very confusing whereas later the road network became clearer.

### V. Conclusion

Water surface as part of ecological system, is an important element of the nature and has environmental qualities for the region, so it has to be treated as an essential component. In this study it was examined the transformation of the coastline of Durrës city, from 'Ura Dajlanit' to 'Plepa'. The analysis of the site is done within a timeframe of 20 years from 1995 to 2015. In this framework the Master Plans, Regulatory Plans, and construction criteria were analyzed and there were found problems with implementation of the buildings in confront with the Urban Law. The transformation of the coastline area from 'Ura Dajlanit' to 'Plepa' below 'Pavarsia' street was observed on the study by using maps, pictures, and changes in Urban Law and Master Plans. From the analysis of the maps during years is seen that there are some informal settlements in this area and an inappropriate use of the land that came as a result of the illegal constructions during years. The massive construction and the problematic of the land use and the buildings typology came out also as result of not respecting the Master Plans and the Urban Laws. There are 2 main factors

which have affected the development of the Coastal area. The first factor is the land reclamation and the second is the absence of an approved Master Plan from 1990 until 2011, since the Regulatory Plan of 1987 was not taken into consideration because it belonged to the previous regime.

Durrës is the second largest city on the region and it has touristic and historical values. The transformation of 'Plazh' area can be seen as a bad example of management the touristic coastal areas. Every city which has an important historical and touristic value should the development plan where all the social, economic, aesthetical and architectural principles should be merged in a harmonic and attractive way. Everyone should enjoy the spaces which are offered by the waterfront areas. This land should be not treated as a personal asset and used for personal interest, but also take into consideration the demands of the tourists and residents. The 'concrete barrier' of the 'Plazh' area was done without a base of a Regulatory Plan or a specific study. Almost 50% of the buildings are done with the violation of the law. At this situation it is difficult to rehabilitate all the area, since there is not any vacant land which can be used for the public spaces. It should be suggested a rehabilitation project, which should include sport areas, recreation areas and more green areas. The new construction should be stopped, some of the informal settlements should be demolished and importance to the secondary roads which orient visitors to the sea should be given.

### References

- [1] A. Hoti, Durrësi : Epidamni - Dyrrahu : guidë, Tiranë: Cetus Tirana, 2003.
- [2] S. Xhafa and B. Hasani, "Urban Planning Challenges in the Peripheral Areas of Durres City (Porto Romano)," *Mediterranean Journal of Social Sciences*, vol. 4(10), 2013.
- [3] R. Leakey and R. Lewin, *People of the Lake: Mankind and Its Beginnings*, Anchor Press/Doubleday, 1978.
- [4] L. Dong, "Waterfront Development: A Case Study of Dalian, China," *University of Waterloo , Waterloo , 2004*.
- [5] M. Moretti, "Cities on Water and Waterfront Regeneration: A Strategic Challenge for the Future," *Rivers of Change - River//Cities*, Warsaw, 2008.
- [6] M. Yassin, B. Azlina, S. Bond and J. McDonagh, "Principles for sustainable riverfront development for Malaysia," *Journal of Techno-Social*, 2012.
- [7] A. Yassin, C. Eves and J. McDonagh, "An Evolution of Waterfront Development in Malaysia," *Geography*, 2010.
- [8] M. Marzia, *Morphological, technological and functional characteristics of infrastructures as a vital sector for the competitiveness of a country system*, Milano: Politecnica, 2011.
- [9] S. Shaziman, I. Usman and M. Tahir, "Waterfront as Public Space Case Study; Klang River between Masjid Jamek and Central Market, Kuala Lumpur," in *Selected topics in energy, environment, sustainable development and landscaping; EEESD '10, 3rd WSEAS International conference on landscape architecture LA '10*, 2010.
- [10] S. Kostof, *The City Shaped: Urban Patterns and Meanings Through History*, Thames & Hudson, 1991.
- [11] N. Erkal, *Haliç extra mural zone : a spatio temporal framework for understanding the architecture of the Istanbul city frontier*, Istanbul: METU, 2001.
- [12] A. R. Al-Shams, K. Ngah, Z. Zakaria, N. Noordin, M. Z. Hilmie and M. Sawal, "Waterfront Development within the Urban Design and Public Space Framework in Malaysia," *Asian Social Science*, vol. 9, no. 10, pp. 77-87, 2013.
- [13] C.-H. Chen, "The Analysis of Sustainable Waterfront Development Strategy - The Case of Keelung Port City," *International Journal of Environmental Protection and Policy*, vol. 3, no. 3, pp. 65-78, 2015.
- [14] Y. Reyhan , Ş. Şenlie and G. İ. Burcu, "Sustainable Urban Design Guidelines For Waterfront Developments," in *2nd International Sustainable Building symposium*, Ankara, 2015.
- [15] R. E. Pramesti, "Sustainable Urban Waterfrontre Development: Challenge And Key Issues," *MEDIA MATRASAIN*, vol. 14, no. 2, pp. 41-54, 2017.
- [16] V. Koçi, *Spatial transformations of the waterfront-as an urban frontier case study : Durres a port city*, Istanbul: METU, 2005.
- [17] Dyrrah, "Durrës: Vilat e plazhit, "hirushet" e vetmuara mes katrahurës ndërtimore," *DurrësLajm*, 15 June 2015.
- [18] G. Enyedi, "Urbanization under Socialism," in *Cities after Socialism, Urban and Regional Change and conflict in Post-Socialist Societies*, Oxford, Blackwell Publishers, 1996, pp. 100-11.

# Deep Learning Using MobileNet for Personal Recognizing

Şafak Kılıç

Department of Computer Engineering  
Siirt University  
Siirt, Turkey  
safakkilic@outlook.com

İman Askerzade

Department of Computer Engineering  
Ankara University  
Ankara, Turkey  
imasker@eng.ankara.edu.tr

Yılmaz Kaya

Department of Computer Engineering  
Siirt University  
Siirt, Turkey  
yilmazkaya1977@gmail.com

**Abstract**— The usage areas of biometric technologies are increasing day by day. As the importance of information security becomes more important for people every day, one of the most used areas has been information security. In recent years, human-computer interactive systems have started to attract academic and commercial interest and it is aimed to solve problems such as person recognition, gender estimation, age estimation with these systems. In our study, person recognition was performed through the data collected using wearable sensors. The Daily and Sports Activities data set, which we obtained from the UCI database, has been tested with the developed MobilNet architecture. It has been seen that the data obtained from the sensors are successful in the person recognition problem. The developed system has realized the person identification with 19 different physical movements and also provided the detection of 19 different movements. In addition, success rates were obtained according to the region where the sensors were installed. Thanks to the results obtained in this study, it has been seen that accelerometer, gyroscope and magnetometric sensors are successful in biometric person recognition. In summary, it has been determined that the proposed method is successful in biometric person recognition, thanks to the data obtained from wearable sensors.

**Keywords**—*Transfer Deep Learning Models, MobileNet, Person Identification, Wearable Sensor, Biometric System*

## I. Introduction

In the past few decades, the problem of identifying people has been one of the hot areas where researchers emphasize the use of various methods. The human body has several unique characteristics. Some systems can detect these characteristics and distinguish them from others. A system that acknowledges people supported their physical or behavioural characteristics is termed a biometric system. Personal biometric authentication includes distinctive people supported their physiological and/or behavioural characteristics [1]. Biometry technology checks the physical or behavioural characteristics that a private will acknowledge. The biometric system works in 2 ways: (1) identification (also referred to as "identity verification") and (2) authentication (also referred to as "identity verification"). First, the verification of private identity is completed by finding a match within the info of everybody within the information (one-to-many comparison strategy). within the latter, a human biometric info is compared with its example hold on within the system information to verify a human identity.

Physical and behavioral features are the two basic qualities of persons. Physical properties are those that are stable and do not vary over time, whereas behavioral properties are those that change over time and in response to environmental factors. Biometric recognition systems are generally developed thanks to these two characteristics of people. While biometric systems created with data such as facial recognition, fingerprint recognition, hand geometry, iris and retina data are systems created by physiological features, biometric systems

created with data such as walking, signature and speech are based on behavioral features. These strategies, on the other hand, have a significant disadvantage in that they can be duplicated [2; 3; 4]

Mimic sounds, the use of duplicate irises, and disguised glasses can be examples of these scams. As a result, new descriptive systems based on individual behavior or features, known as biometrics, based on signals measured from various parts of the body, have been adopted in recent years [5; 6]. Different medical signals are also employed as biometric data, according to studies. Biometric systems are developed using EEG and other signals [7; 8; 6; 9], electrocardiogram [10; 11; 12; 13; 14; 15], and accelerometer [16; 17]. Medical indicators are unique to each person, according to studies [8; 18; 10].

In the study of Alyasseri et al. [19], people are recognized using multichannel EEG waves. In addition, active EEG channels were uncovered by the researchers. The process of recognizing persons is done by the use of electrical signals in the brain, according to Sun et al. [9]. They discovered that applying the conventional 1D-LSTM deep learning algorithm to 16-channel EEG measurements resulted in a success rate of 99.56 percent. In their study of identifying people using EEG data, Rodrigues et al. discovered an 87 percent success rate [6].

The person recognition problem was attempted to be solved by using sensor signals in a study by Kılıç et al. [20; 21]. The sensor signals were converted into pictures with various processes and the success of the system was tested with local binary pattern and deep learning networks and models. In both of his articles, the author showed a success rate of over 95% in the person recognition study from sensor tests.

Although the CNN training strategy from scratch can be successful in many problems, the correct optimization of the hyper-parameters in the architecture to be installed is still a difficult process [34]. At the same time, a large amount of data is required for a scratch-training technique [35]. However, it is possible to reach high success rates faster and more precisely by training deep architectures, which are enriched with techniques newly introduced to the literature and also do not need hyperparameter optimization, with transfer learning or fine-tuning strategy [36]. The MobileNet architecture is chosen as the training model in this study because it is learned using deep transfer networks, has a low transaction cost, and is ideal for mobile applications. At the end of the study, higher success rates were achieved with the right education strategy compared to other studies in the literature.

In our study, after this stage, the data set was introduced, the models used were shown, and the experimental results obtained were presented in the discussion section. In the last part, the general achievements obtained at the end of my study are given.



## II. Data Set

As a part of this study, dataset which is known as a Daily and Sports Activities obtained from the ICU database [22; 23; 24]. This study used Xsens MTx sensors mounted to the person's assigned location to collect data from the 19 previously indicated behaviors (activities). To collect data, place these sensors in 5 different areas of the topic. The chest, right wrist joint, left wrist, right (above knee), and left leg (above knee) have all been identified as possible locations for the device (Figur1). In each Xsens MTx device, there are nine sensors (accelerometer x, y, z; gyroscope x, y, z and magnetometer x, y, z).

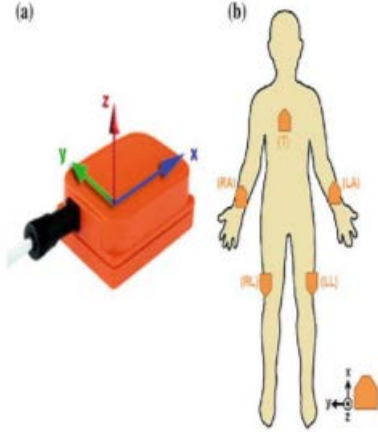


Figure I. Attaching the sensors to 5 different regions of the subjects. (A) Xsens MTx, (B) Regions where the sensors are attached [33].

The data for this study was generated by four women and four men participating in nineteen distinct planned activities for five minutes each. Table I lists the activities carried out by the topics.

Table I. The study involved 19 different activities.

Activity Code	Name of The Activity
A1	the act of sitting
A2	the act of standing
A3	recumbency exercise
A4	Lie down on your right side
A5	transferring to the second floor
A6	descending to the lower level
A7	in the elevator, standing activities
A8	while the elevator is going, standing activities
A9	In the parking lot, there is a lot of walking going on.
A10	Walking exercise on a treadmill at a speed of 4km/h parallel to the ground
A11	Walking on the treadmill at a 15-degree inclination to the ground at a speed of 4 km/h
A12	Running at a speed of 8 km/h is a high-intensity activity.
A13	Running at an 8-kilometer-per-hour pace is a high-intensity activity.
A14	riding on an elliptical machine
A15	Cycling in a horizontal position is an activity that involves cycling in a horizontal position.

Activity Code	Name of The Activity
A16	Vertical cycling is an activity that involves cycling in a vertical position.
A17	rowing exercise
A18	jumping exercise
A19	basketball game in progress

## III. Methodology

### A. Person Identification By MobileNet Deep Transfer Learning Technique

The figure below shows the recommended method of identifying people using the MobileNet deep learning method. A brief introduction to the process performed at each stage.

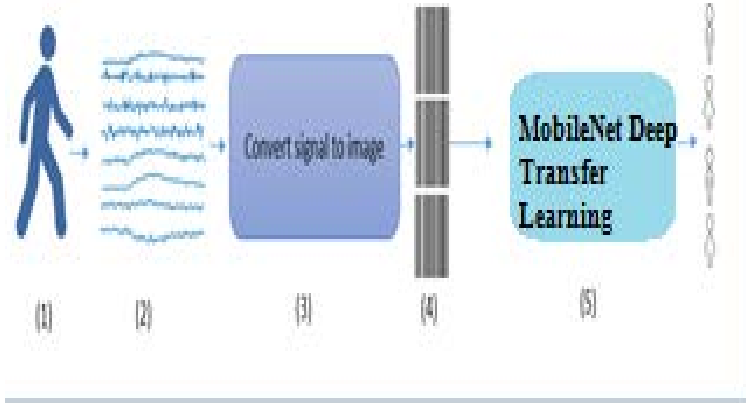


Figure II MobileNet deep transfer learning approach template for person identification (PI-MobileNet )

Block 1: A wearable sensor is attached to the subject at this point. The Xsens MTx sensor devices are placed throughout the subject in five separate locations.

Block 2: A total of 45 signal channels from 5 different places of the body are recorded using accelerometer, gyroscope, and magnetometer sensors.

Block 3: Each action in Table I has a 5-minute signal measurement. These signals are then separated into 5 seconds. The signal length is  $5 \times 25 = 125$  because the sampling rate is 25. First, use the following equation to convert the value of the signal to a value between 0-255. Since the number of channels is 45, you get an image in the form of  $125 \times 45$ .

$$New X_i = round \left( \left[ \frac{X_i - Min(X)}{Max(X) - Min(X)} \right] \times 255 \right) \quad (1)$$

Figure III, for example, depicts the image formed by the standing activity signal for each individual.

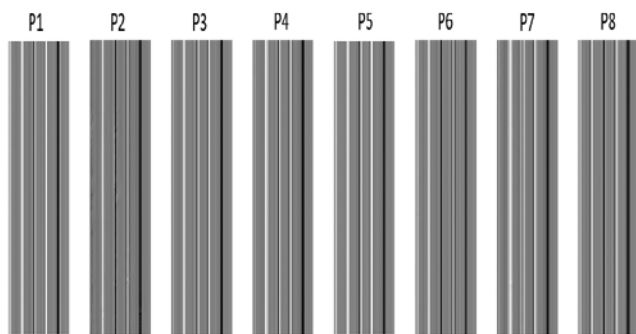


Figure III. For each subject, images of the signals of sitting activity.

Block 4: The image that was made is shown. These photos are sent to the Deep Transfer Learning algorithm for restoration.

Block 5: The image is classified by MobileNet depth transfer learning technology. All architectures of version 1 and 2 of MobileNet have been tested.

### B. Deep Transfer Learning

AlexNet [25], GoogLeNet [26], VGG Net [27], ResNet [28], and NASNet [29] are just a few of the CNN architectures that have been developed in recent years. All of these architectures have been pre-trained on over one million photos from ImageNet's Large-Scale Visual Recognition Challenge [30]. To transmit these learnings, we can use the model's weights and biases. One of the pre-trained models is the MobileNet architecture [31]. The distinction between MobileNet and other published designs is that MobileNet has a lesser number of parameters. With separable convolution, MobileNet exploits the depth introduced in to lower the computational requirements of the first layers [32]. Latency and precision are trade-offs in MobileNet efficiency. Tasks based on this model can be completed on the CPU without the need of the GPU due to MobileNet's light weight. They, like other large-scale models, can be utilized for classification, detection, and segmentation.

### C. The Architecture Of MobileNet

Deep learning networks produce effective outcomes in areas like image analysis and computer vision [5]. There are many hidden layers in a DL. The values given as input to these layers form outputs by going through mathematical operations. These outputs are given as input to the next layer (Figure-1). DL is trained on the designated classes for image analysis and classification. This training requires a previously labeled dataset that DL will work on. Because the analysis of images requires high computational power, configuring the system in which the DL solution will run is very important for performance.

MobileNet is built on a separable depth convolution with two basic layers: depth and point convolution. Deep convolution is the process of filtering the input without adding additional functions. As a result, point-by-point convolution is used to combine the process of creating new features. Finally, a depth separable convolution is the result of combining the two layers. The model applies a single filter

to each input channel using deep convolution and then creates a linear combination of deep layer outputs using 1x1 convolutions (point by point). Batch normalisation (BN) and changed long measure (ReLU) liner unit used once every convolution.

Looking at Table II, the MobileNet architecture is seen in detail. The picture given to the model is transformed into the output image by going through the stages in the table. ReLu was chosen as the activation function. This model has approximately 4 million parameters, which is significantly less than the other models.

Table II. The Architecture Of MOBILENET

MobileNet
Input layer
Convolutional layer
Depthwise n Convolution layer Batch Normalization ReLU + Pointwise n Convolution layer Batch Normalization ReLU (n = 1, 2, 3, ..., 13 layers)
Global Average Pooling layer
Reshape layer
Dropout layer
Convolutional layer
Softmax layer
Reshape layer
Output

### IV. Experiment

In this study, we use a lightweight network called MobileNet to develop a deep neural network architecture that uses a separable deep convolution.

Although the design employs fewer parameters than previous efficiency models at the same level, Google designed MobileNet with the goal of striking a balance between latency and precision and displaying impressive performance. We propose our upgraded MobileNet, which is appropriate for personal recognition categorization utilizing wearable sensor inputs. Starting with the primary convolutional layer, followed by thirteen deep convolutional layers, and at last a degree convolutional layer, our design is galvanized by the MobileNet network layer. once every depth and purpose convolutional layer, use the batch standardization (BN) and changed linear measure (ReLU) trigger functions. the worldwide average grouping layer are accustomed minimize the dimensions of the extracted feature

map once all of the convolutional layers have extracted options from the input image.

The Reshape layer, Dropout layer, convolutional layer, Softmax activation perform layer, and thus the Reshape layer, which make up the last 5 layers of the quality MobileNet, are replaced by the Dropout layer and therefore the completely connected layer, which uses the Softmax activation perform. Our fully linked layer will generate more accurate predictions for each class than the preceding five layers using standard MobileNet. Increasing the number of convolutional layers in the model will help it extract more alternatives from the computer file in general.

However, there are some limitations to adding them. We discovered in this work that traditional MobileNet overfits, causing the model to misclassify personel recognition. As a result, our proposed design is capable of resolving this issue. Our modified model's total parameters are reduced, which reduces the time it takes to calculate the model. Table III shows the specifics of our architecture.

Table III. Our MobilNet Articheture

<b>MobileNet</b>
Input layer
Convolutional layer
Depthwise n Convolution layer Batch Normalization ReLU +
Pointwise n Convolution layer Batch Normalization ReLU (n = 1, 2, 3, ..., 13 layers)
Global Average Pooling layer
<b>Dropout layer</b>
<b>Fully Connected layer</b>
Output

MobileNet is a pre-trained model that recognizes the shape and part of an object in its initial layer and was trained using the ImageNet dataset. So in this paper, we chose all layer deep transfer learning for traditional MobileNet and compared it to our modified MobileNet, which is only two-thirds of the layer ratio and the remaining layers configure the Learn migration. We used our data set and settings to train the model (such as weights, bias, and learning rate). Finally, we assessed MobileNet and our model using the original data set from the previous experiment, without any data augmentation or sampling techniques. MobileNet and our model are tested in the second experiment utilizing data that has been completed by sampling and augmenting data.

## V. Results

Signals from 19 different activities of 8 people were used to create the data set for this investigation (4 men and 4

women). Each activity is broken down into 60 segments. As a result, there are  $19 \times 8 \times 60 = 9120$  signal matrices in the data set. MobileNet deep transfer learning technology is utilized after these signal matrices have been converted into images. 9120 photos were extracted as a consequence of to see if our system worked. There are two MobileNet architectures in operation. The success rate is determined as follows:

$$100 * \frac{\# \text{ True classified}}{\# \text{ True classified} + \# \text{ False classified}} (\%) \quad (2)$$

Table IV shows the success rate of personnel identification using MobileNet Version 1 (MobileNetV1) and MobileNet Vesiyon 2 (MobileNetV2) architectures.

Table IV. Person Recognition Success Of mobilet Networks

<b>Model</b>	<b>Success rate%</b>
MobileNetV1	96,82
MobileNetV2	<b>98,46</b>

It is seen in Table IV that the person recognition problem with the data obtained from the sensor signals is successfully solved with MobileNet networks. Although success rates are close to each other, MobilnetV2 networks were more successful in person recognition problem than MobileNetv1 networks.

Table V. Success Rate of Mobilenet networks by Physical Activity

<b>Activity</b>	<b>MobileNetV1</b>	<b>MobileNetV2</b>
A1	100	64,5
A2	100	12,5
A3	97,9	25
A4	93,75	47,916
A5	93,7	18,7
A6	95,8	22,9
A7	97,9	6,2
A8	62,5	16,6
A9	95,8	4,1
A10	100	10,4
A11	100	16,6
A12	100	16,6
A13	95,8	95,8
A14	97,9	22,9
A15	91,6	18,7
A16	95,8	16,6
A17	87,5	27
A18	79,1	27
A19	75	56,2

We also tried our model from the physical activity recognition problem to measure the success of our network

after the person recognition process. In general, it has been determined that the MobileNet V1 network is more successful than the physical activation recognition problem. It has shown 100% success in recognizing some physical movements. Detailed results are presented in Table V.

Table VI. Success Rate of Mobilenet networks by region of sensors

Region	MobileNetV1	MobileNetV1
Chest	94,5	93,8
Right Arm	92,1	92,1
Left Arm	91,1	89,9
Right Leg	93,6	92,3
Left Leg	91,9	92

Finally, we have tested the success of our model according to the region where the sensors receive signals. Both models showed high success. In addition, MobileNetv1 networks have a higher success rate. Data from chest-level sensors was more distinctive than data from other regions. Details can be seen in Table VI.

## VI. Conclusion

Several biometric technologies have been developed in recent years. Face, voice, fingerprints, palm print, ear shape, and gait are all biometric technologies that have been widely used in security systems. However, because they may be imitated, most of these systems have glaring faults. To address these issues, a new biometric system based on medical signals has been developed. A biometric approach was developed in this study to identify people using wearable sensor inputs. The major goal of this study is to show that signals from portable sensors such as accelerometers, gyroscopes, and magnetometers may be used to identify people. In the future studies, due to the increasing use of mobile devices, person recognition will be done with a deep transfer learning approach through data obtained from portable mobile devices.

## References

- [1] B. Ngugi, A. Kamis, and M. Tremaine, "Intention to use biometric systems," *e-Service Journal: A Journal of Electronic Services in the Public and Private Sectors*, vol. 7, no. 3, pp. 20-46, 2011.
- [2] T. Matsumoto, H. Matsumoto, K. Yamada, and S. Hoshino, "Impact of artificial "gummy" fingers on fingerprint systems," in *Optical Security and Counterfeit Deterrence Techniques IV*, 2002, vol. 4677: International Society for Optics and Photonics, pp. 275-289.
- [3] J. Yu, C. Fang, J. Xu, E.-C. Chang, and Z. Li, "ID repetition in Kad," in *2009 IEEE Ninth International Conference on Peer-to-Peer Computing*, 2009: IEEE, pp. 111-120.
- [4] G. Gainotti, "Laterality effects in normal subjects' recognition of familiar faces, voices and names. Perceptual and representational components," *Neuropsychologia*, vol. 51, no. 7, pp. 1151-1160, 2013.
- [5] J. Galbally, S. Marcel, and J. Fierrez, "Biometric antispoofing methods: A survey in face recognition," *IEEE Access*, vol. 2, pp. 1530-1552, 2014.
- [6] D. Rodrigues, G. F. Silva, J. P. Papa, A. N. Marana, and X.-S. Yang, "EEG-based person identification through binary flower pollination algorithm," *Expert Systems with Applications*, vol. 62, pp. 81-90, 2016.
- [7] S. Marcel and J. d. R. Millán, "Person authentication using brainwaves (EEG) and maximum a posteriori model adaptation," *IEEE transactions on pattern analysis and machine intelligence*, vol. 29, no. 4, pp. 743-752, 2007.
- [8] Y. Dai, X. Wang, X. Li, and Y. Tan, "Sparse EEG compressive sensing for web-enabled person identification," *Measurement*, vol. 74, pp. 11-20, 2015.
- [9] Y. Sun, F. P.-W. Lo, and B. Lo, "EEG-based user identification system using 1D-convolutional long short-term memory neural networks," *Expert Systems with Applications*, vol. 125, pp. 259-267, 2019.
- [10] S. A. Israel, J. M. Irvine, A. Cheng, M. D. Wiederhold, and B. K. Wiederhold, "ECG to identify individuals," *Pattern Recognition*, vol. 38, no. 1, pp. 133-142, 2005.
- [11] M. Deng, C. Wang, M. Tang, and T. Zheng, "Extracting cardiac dynamics within ECG signal for human identification and cardiovascular diseases classification," *Neural Networks*, vol. 100, pp. 70-83, 2018.
- [12] A. Goshvarpour and A. Goshvarpour, "Human identification using a new matching pursuit-based feature set of ECG," *Computer methods and programs in biomedicine*, vol. 172, pp. 87-94, 2019.
- [13] K. Su et al., "Human identification using finger vein and ECG signals," *Neurocomputing*, vol. 332, pp. 111-118, 2019.
- [14] F. Sufi and I. Khalil, "Faster person identification using compressed ECG in time critical wireless telecardiology applications," *Journal of Network and Computer Applications*, vol. 34, no. 1, pp. 282-293, 2011.
- [15] W. Chang, H. Wang, G. Yan, and C. Liu, "An EEG based familiar and unfamiliar person identification and classification system using feature extraction and directed functional brain network," *Expert Systems with Applications*, vol. 158, p. 113448, 2020.
- [16] R. San-Segundo, R. Cordoba, J. Ferreiros, and L. F. D'Haro-Enriquez, "Frequency features and GMM-UBM approach for gait-based person identification using smartphone inertial signals," *Pattern Recognition Letters*, vol. 73, pp. 60-67, 2016.
- [17] R. San-Segundo, J. D. Echeverry-Correa, C. Salamea-Palacios, S. L. Lutfi, and J. M. Pardo, "I-vector analysis for gait-based person identification using smartphone inertial signals," *Pervasive and Mobile Computing*, vol. 38, pp. 140-153, 2017.
- [18] N. V. Boulgouris, K. N. Plataniotis, and E. Micheli-Tzanakou, *Biometrics: theory, methods, and applications*. John Wiley & Sons, 2009.
- [19] Z. A. A. Alyasseri, A. T. Khader, M. A. Al-Betar, and O. A. Alomari, "Person identification using EEG channel selection with hybrid flower pollination algorithm," *Pattern Recognition*, vol. 105, p. 107393, 2020.
- [20] Ş. Kılıç, Y. Kaya, and I. Askerbeyli, "A New Approach for Human Recognition Through Wearable Sensor Signals," *Arabian Journal for Science and Engineering*, vol. 46, no. 4, pp. 4175-4189, 2021.
- [21] Ş. Kılıç, İ. Askerzade, and Y. Kaya, "Using ResNet Transfer Deep Learning Methods in Person Identification According to Physical Actions," *IEEE Access*, vol. 8, pp. 220364-220373, 2020.
- [22] K. Altun, B. Barshan, and O. Tunçel, "Comparative study on classifying human activities with miniature inertial and magnetic sensors," *Pattern Recognition*, vol. 43, no. 10, pp. 3605-3620, 2010.
- [23] B. Barshan and M. C. Yükek, "Recognizing daily and sports activities in two open source machine learning environments using body-worn sensor units," *The Computer Journal*, vol. 57, no. 11, pp. 1649-1667, 2014.

- [24] K. Altun and B. Barshan, "Human activity recognition using inertial/magnetic sensor units," in International workshop on human behavior understanding, 2010: Springer, pp. 38-51.
- [25] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Advances in neural information processing systems*, vol. 25, pp. 1097-1105, 2012.
- [26] C. Szegedy et al., "Going deeper with convolutions," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2015, pp. 1-9.
- [27] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," arXiv preprint arXiv:1409.1556, 2014.
- [28] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 770-778.
- [29] B. Zoph, V. Vasudevan, J. Shlens, and Q. V. Le, "Learning transferable architectures for scalable image recognition," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2018, pp. 8697-8710.
- [30] O. Russakovsky et al., "Imagenet large scale visual recognition challenge," *International journal of computer vision*, vol. 115, no. 3, pp. 211-252, 2015.
- [31] A. G. Howard et al., "Mobilenets: Efficient convolutional neural networks for mobile vision applications," arXiv preprint arXiv:1704.04861, 2017.
- [32] L. Sifre and S. Mallat, "Rigid-Motion Scattering for Image Classification. arXiv 2014," arXiv preprint arXiv:1403.1687.
- [33] O. Dobrucali, B. Barshan. (2013) Sensor-Activity Relevance in Human Activity Recognition with Wearable Motion Sensors and Mutual Information Criterion. In: Gelenbe E., Lent R. (eds) Information Sciences and Systems 2013.
- [34] M. Feurer and F. Hutter, "Hyperparameter optimization," in Automated machine learning: Springer, Cham, 2019, pp. 3-
- [35] D. Shen, G. Wu, and H.-I. Suk, "Deep learning in medical image analysis," *Annual review of biomedical engineering*, vol. 19, pp. 221-248, 2017.
- [36] D. Hendrycks, K. Lee, and M. Mazeika, "Using pre-training can improve model robustness and uncertainty," in International Conference on Machine Learning, 2019: PMLR, pp. 2712-2721.

# Effects of Photon-Shot and Excess Noises on Detectable Minimum Rotation Rate in I-FOG Design for Autonomous Vehicles

Kübra Kılınçarslan  
Department of Electrical and Electronic  
Engineering  
Bursa Uludag University  
Bursa, Turkey  
kubrakilincarslan@uludag.edu.tr

Emirhan Sağ  
Department of Electrical and Electronic  
Engineering  
Bursa Uludag University  
Bursa, Turkey  
emirhansag@uludag.edu.tr

Abdurrahman Günday  
Department of Electrical and Electronic  
Engineering  
Bursa Uludag University  
Bursa, Turkey  
agunday@uludag.edu.tr

**Abstract**—In the last decade, the importance of acquiring location information is gradually increased depending on the developments in autonomous vehicle technologies. The rotation rate variation required for providing high precision continuous position information is obtained by using gyroscope, which is an important part of the Inertial Measurement Unit (IMU). However, some effects limit the operation performance of gyroscopes. In this study, the effects of photon shot and excess noises that limit the measurement accuracy of interferometric fiber optic gyroscope (I-FOG) on detectable minimum rotation rate (DMRR) have been analyzed for a basic configuration using Superluminescence diode (SLD) and Superfluorescent fiber source (SFS). Furthermore, simulations related to DMRR variations for the system parameters of fiber length, fiber coil diameter, photodetector bandwidth, output power of the optical sources and spectral bandwidth have been obtained in Matlab 2020b. Moreover, for an optimum system design employing SLD and SFS, DMRR has also been computed as 0.793°/h and 0.910°/h, respectively. Thus, approximately 80% improvement has been achieved over the DMRR values in the system.

**Keywords**— fiber optic gyroscope, broadband sources, minimum rotation rate, photon-shot noise, excess noise

## I. Introduction

Acceleration sensors and gyroscopes with high sensitivity and accuracy are extensively used to detect the position, speed and direction of an object. These kinds of information gathered from inertial navigation systems (INS) are combined with global positioning system (GPS) data, accordingly [1].

In recent years, many research and development studies have been carried out in the field of autonomous driving technologies. In this context, detection systems such as GPS, Visual Simultaneous Localization and Mapping (SLAM), and Light Detection and Ranging (LIDAR) are used to obtain the exact location information of the vehicle [2].

However, these methods and techniques have some limitations in terms of performance due to their weaknesses in practical studies. Some of these restrictions are atmospheric effects, multi-path propagation and inaccessibility to the mediums of tunnels for GPS, affecting adversely due to the atmospheric events for image processing in Visual SLAM, exhibiting capability of different performance for all kinds of objects depending on the reflection of light for LIDAR systems.

In cases where other positioning methods cannot be used, the dead reckoning technique is utilized to overcome the weaknesses of these positioning methods. Dead reckoning is in principle the process of calculating the current position of a person or moving object by using the fixed or previously specified position and advancing this position utilizing known

or estimated information of vehicle speed and route. There are great numbers of sensors used in the computational navigation system with high accuracy and sensitivity. The gyroscope that is one of them is exploited for detecting the deviation rate and route information and becomes the most important device that determines the accuracy of the dead reckoning [3].

There are different types of gyroscopes employed in autonomous vehicles such as optical-based and micro-electro-mechanical systems (MEMS). MEMS gyroscopes have cost advantages in comparison with optical gyroscopes, on the contrary, optical gyroscope-based navigation systems are preferred due to their superiorities in terms of measurement accuracy and reliability [4].

However, optical backscattering and distortions caused by the nonlinear electro-optical effects influence the operation performances of I-FOGs [5, 6]. To overcome this problem, low coherent and broadband optical sources are employed in applications [7]. The simulations and analysis performed in this study have been obtained making use of SLD and SFS that are low coherent and broadband optical sources.

The main effect limiting the DMRR at the optical gyroscope output is the photon shot noise induced by the photodetector that converts the incoming optic signal into an electrical signal in the FOG configuration exploited in this study. Furthermore, excess noise induced by characteristics of optical sources [6] is considered as a limiting effect on DMRR.

In this study, performance values of DMRR for I-FOGs designs employing SLD and Erbium-doped SFS have been compared with each other using theoretical and the simulation results derived from the configuration illustrated in Figure I. Moreover, DMRR variations and corresponding simulations have also been performed making use of the parameters of optical fiber coil diameter, fiber length, photodetector bandwidth and optical output power and spectral bandwidth of both sources.

## II. Theory

The working principle of optical gyroscopes is based on the Sagnac effect. This effect gives rise to the  $\Delta\phi$  Sagnac phase shift between optical signals propagating clockwise (CW) and counterclockwise (CCW) in a ring interferometer rotating around an axis perpendicular to the ring. Detectable minimum rotation rate information is obtained by using the Sagnac phase shift.

The minimum FOG configuration exploited in this study is shown in Figure I [8]. This configuration consists of an optical source, 50/50 optical coupler, fiber polarizer, piezoelectric phase modulator, fiber coil and photodetector.

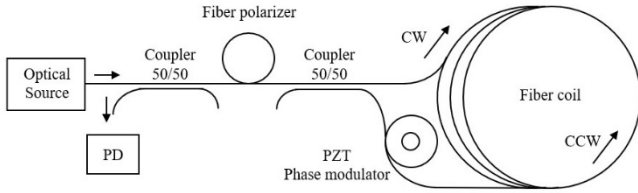


Figure I. Minimum I-FOG configuration

The fiber polarizer used in this FOG illustration given in Figure I. polarizes the light linearly in only one direction. A 50/50 optical coupler placed in the configuration is utilized for splitting the light beam into two equal parts propagating in opposite directions of each other in the fiber coil. An optical phase modulator is utilized to modulate the phase of the light pumped into the optical fiber. The phase difference variation caused by the rotational movement into the system reaches towards the photodetector and then it is converted into an electrical signal [9]. Afterward, rotation rate information is obtained by using the signal processing unit at the photodetector output.

The relation between the rotation rate caused by the rotational motion in I-FOGs and the Sagnac phase shift is given as in (1).

$$\Omega = \frac{\lambda \cdot c}{2 \cdot \pi \cdot L \cdot D} \cdot \phi_s \quad [\text{rad/s}] \quad (1)$$

where  $\Omega$  is the rotation rate,  $\lambda$  is the wavelength,  $c$  is the speed of light,  $L$  is the fiber length,  $D$  is the diameter of the coil and  $\phi_s$  is Sagnac phase shift.

In I-FOGs, the fundamental noise limit is described by photon-shot noise, which causes random fluctuations in the detector output current ( $i_s = \sqrt{(q_i B)}$ ) due to the random scattering of incident photons on the photodetector.

In the noise current equation, the parameters of  $q$ ,  $i_0$ ,  $B = 1/T$  and  $T$  represent the electron charge, the average current in the detector, the electrical bandwidth of the detection system and the sampling time, respectively [10].

I-FOG resolution for the gyroscope system is evaluated as the detectable minimum variation rate in rotation angle induced by uncertainties in detector output current and rotation rate. The detectable minimum rotation rate for photon shot noise is expressed by (2) [11].

$$\Omega_{\min, \text{shot}} = \frac{\lambda \cdot c}{2 \cdot \pi \cdot L \cdot D} \sqrt{\frac{B \cdot h \cdot c}{\eta \cdot \lambda \cdot P_d}} \quad [\text{rad/s}] \quad (2)$$

where  $B$  is the bandwidth of the photodetector,  $h$  is the Planck constant,  $\eta$  is the optical efficiency of the detector, and  $P_d$  is the optical power to the detector. Photodetector current is stated with (3) [12].

$$I = \frac{\eta \cdot \lambda \cdot P_d}{h \cdot c} q \quad [\text{A}] \quad (3)$$

where  $I$  is the photodetector current in Ampere. Substituting (3) in (2) and reorganizing the equation, detectable minimum rotation rate as a function of photodetector current can be written as in (4).

$$\Omega_{\min, \text{shot}} = \frac{\lambda \cdot c}{2 \cdot \pi \cdot L \cdot D} \sqrt{\frac{B \cdot q}{I}} \quad [\text{rad/s}] \quad (4)$$

Detectable minimum rotation rate depending on excess noise can be stated given as in (5) [13].

$$\Omega_{\min, \text{excess}} = \frac{\lambda^2}{2 \cdot \pi \cdot L \cdot D} \sqrt{\frac{B \cdot c}{\Delta \lambda}} \quad [\text{rad/s}] \quad (5)$$

where  $\Delta \lambda$  is the spectral bandwidth of the optical source.

Photodetector current generated in the photodetector can be described as a function of the combined effect of photon shot noise and excess noise as given in (6) [14].

$$\langle (\Delta I)^2 \rangle = 2q(I)B + \frac{(I)^2 B}{\Delta \lambda} \quad [\text{A}] \quad (6)$$

Considering the combined effect of photon shot and excess noise, i.e. using (4) and (5), the detectable minimum rotation rate can be expressed as (7) [14].

$$\Omega_{\min} = \frac{\lambda \cdot c}{2 \cdot \pi \cdot D \cdot L} \left( \frac{1+J_0}{J_1} \right) \left( \frac{q}{(I)} + \frac{\lambda^2}{2 \cdot c \cdot \Delta \lambda} \right)^{\frac{1}{2}} \sqrt{B} \quad [\text{rad/s}] \quad (7)$$

where  $J_0$  and  $J_1$  are Bessel functions and take the values of 0.34 and 0.58, respectively for maximum sensitivity.

### III. Simulations

Simulations have been performed using Matlab 2020b on the simulation model built up for analyzing the effects of photon shot and excess noises on the DMRR utilizing the performance parameters of fiber coil diameter, fiber length, photodetector bandwidth, optical source bandwidths, and optical output power. The photon shot noise, excess noise and combined effect of both noises expressed in (4), (5) and (7), respectively, have been exploited for obtaining relevant simulations in this research.

Thereafter, system parameters used in I-FOG configuration have been specified and their values have also been acquired for an optimum system design. In this way, the effects of design parameters on the noise characteristics have been analyzed in terms of system performance, as well.

Parameters of the system used for getting simulations in this study and their corresponding values are given as in Table I.

Table I. Simulation parameters

Parameters	Value ranges
Fiber coil diameter (D) [m]	0.05 – 0.15
Fiber length (L) [m]	1000 – 1500
Photodetector bandwidth (B) [Hz]	60 – 500
SLD spectral bandwidth ( $\Delta \lambda$ ) [nm]	70 – 130
SFS spectral bandwidth ( $\Delta \lambda$ ) [nm]	20 – 60
Output power (P) [mW]	2 – 20

For the I-FOG configuration indicated in Figure I, the wavelength of the light launched into the optical fiber ( $\lambda$ ), electron charge ( $q$ ), speed of light in vacuum ( $c$ ) and photodetector responsivity values have been taken as 1550 nm,  $1.602 \times 10^{-19}$  C,  $3 \times 10^8$  m/s and 0.93 A/W, respectively in the simulations.

### A. Relationship between Minimum Rotation Rate and D.L product

Spectral bandwidths of SLD and SFS vary in ranges of 70 nm - 130 nm and 20 nm - 60 nm, respectively as mentioned in Table I. Simulations with respect to the effects of fiber coil diameter and fiber length products (D.L) on minimum rotation rate variations have been obtained for values of B is 60 Hz, P is 2 mW and spectral bandwidths of SLD and SFS are 70 nm and 20 nm, respectively, as shown in Figure II.

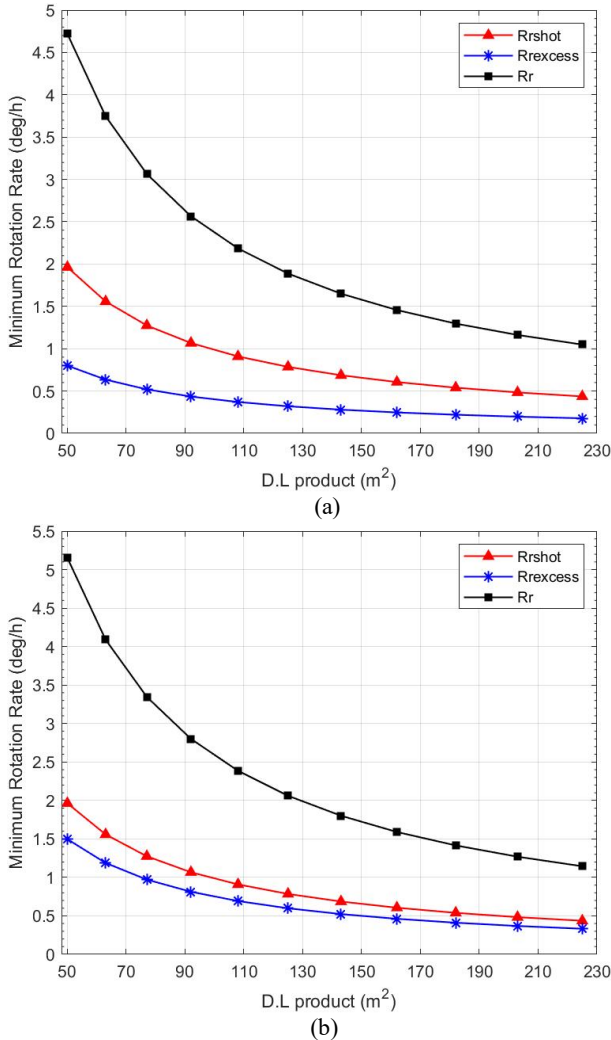


Figure II. Minimum rotation rate vs. D.L product for a) SLD and b) SFS

DMRR variations for both SLD and SFS decrease exponentially with increasing the products of D.L as shown in Figure II. For the variations of D.L in the range of 50 m<sup>2</sup> - 225 m<sup>2</sup>, whilst DMRR changes between 1.963°/h and 0.436°/h depending on the photon shot noise for SLD, their values vary between 0.799°/h and 0.178°/h depending on the excess noise.

In case the variations of D.L are in the same range, DMRR changes between 1.963°/h and 0.436°/h depending on the photon shot noise whilst it changes between 1.496°/h and 0.333°/h depending on excess noise for SFS.

The total DMRR shows a change in ranges of 4.720°/h - 1.049°/h and 5.152°/h - 1.145°/h for SLD and SFS, respectively. Therefore, it is seen that DMRR decreases as the values of coil diameter and fiber length increase.

At the point where the D.L is 150 m<sup>2</sup> in Figure II, values of DMRR for SLD and SFS are decreased 85%, approximately. The detectable minimum rotation rate gets the values of 1.573°/h and 1.717°/h for SLD and SFS, respectively at this point.

### B. Relationship between Minimum Rotation Rate and Bandwidth

Simulations related to the effects of photodetector bandwidth changes on the DMRR variations for DL is 150 m<sup>2</sup> and P is 2 mW have been attained for SLD and SFS with spectral bandwidths of 70 nm and 20 nm as indicated in Figure III, respectively.

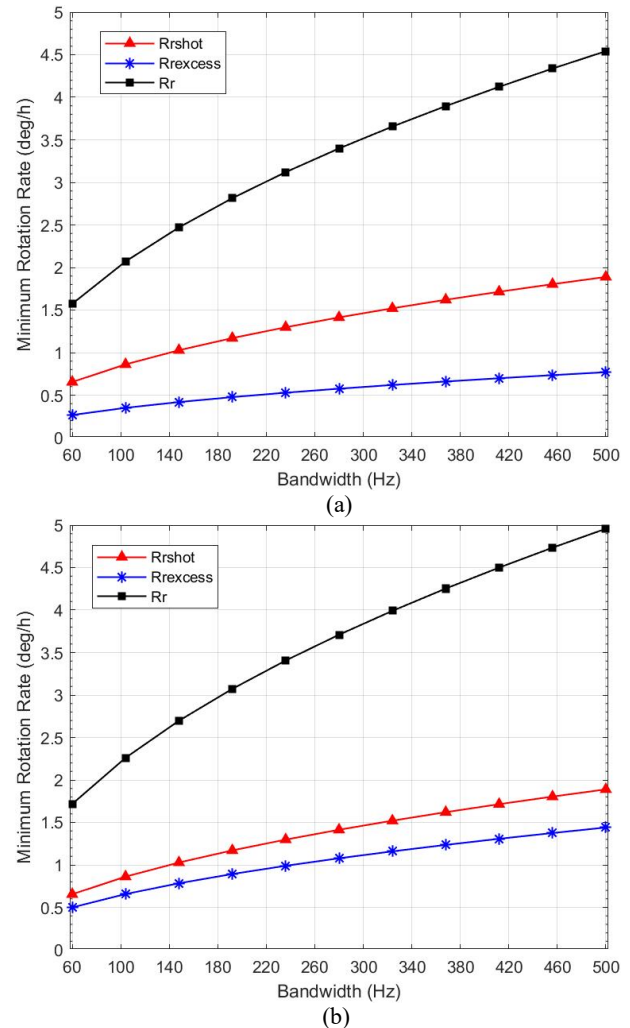


Figure III. Minimum rotation rate vs bandwidth for a) SLD and b) SFS

For the variation of bandwidth in the range of 60 Hz - 500 Hz, DMRR values for SLD varies in the range of 0.654°/h - 1.889°/h depending on the photon shot noise and varies in the range of 0.266°/h - 0.769°/h depending on the excess noise. In case the variation of bandwidth in the same range, DMRR shows a change from 0.654°/h to 1.889°/h depending on the photon shot noise and from 0.499°/h to 1.440°/h depending on the excess noise for SFS.

The total DMRR changes in ranges of 1.573°/h - 4.541°/h and 1.717°/h - 4.958°/h for SLD and SFS, respectively. In other words, the higher the photodetector bandwidth is, the higher DMRR is. The minimum rotation rate reaches its maximum value with 4.958°/h because of the combined effect of both the photon shot noise and the excess noise for SFS.



### C. Relationship between Minimum Rotation Rate and Spectral Bandwidth

The variations of DMRR with the spectral bandwidths of the optical sources have been obtained as shown in Figure IV, when D.L = 150 m<sup>2</sup>, P = 2 mW, and B = 60 Hz.

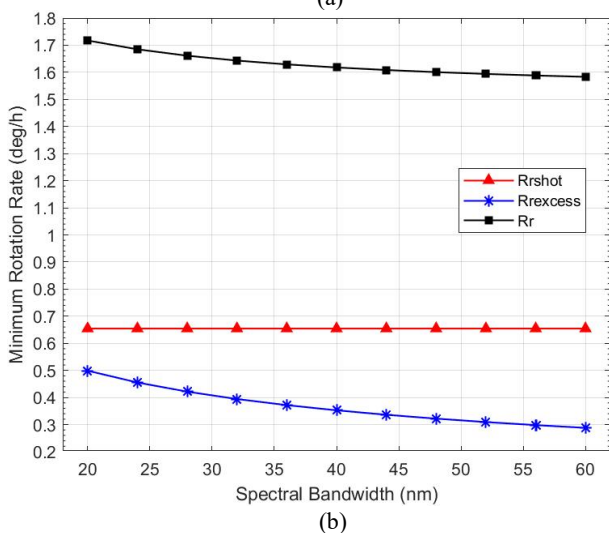
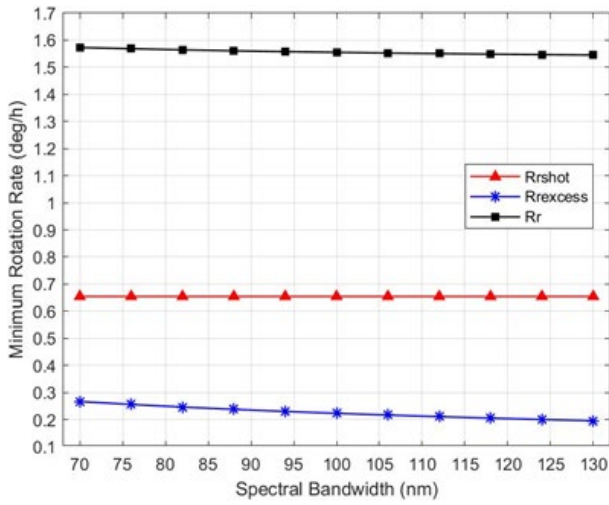


Figure IV. Minimum rotation rate vs spectral bandwidth for a) SLD and b) SFS

The total minimum rotation rate shows a change from 1.573°/h to 1.545°/h for SLD spectral bandwidth variation in the range of 70 nm - 130 nm whilst they vary from 1.717°/h to 1.583°/h for SFS spectral bandwidth variation in the range of 20 nm - 60 nm.

The total DMRR has been obtained 1.555°/h and 1.618°/h for SLD and SFS with the spectral bandwidths of 100 nm and 40 nm, respectively. Hence, the improvements with the values of 64 % and 74 % have been achieved at these points, consecutively.

As it is obvious in Figure IV.b, excess noise is more effective in SFS with narrow bandwidth in comparison to the one with broadband. Stated in other words, this noise effect causes the detectable total minimum rotation rate to become higher in SFSs. In this manner, when using a broadband optical source, the effect of excess noise can be reduced and a lower rotation rate can be achieved.

### D. Relationship between Minimum Rotation Rate and Optical Output Power

Simulations relevant to the effects of the output powers on DMRR for the spectral bandwidth of 70 nm and 20 nm for SLD and SFS, respectively and with D.L = 150 m<sup>2</sup>, B = 60 Hz have been attained as illustrated in Figure V.

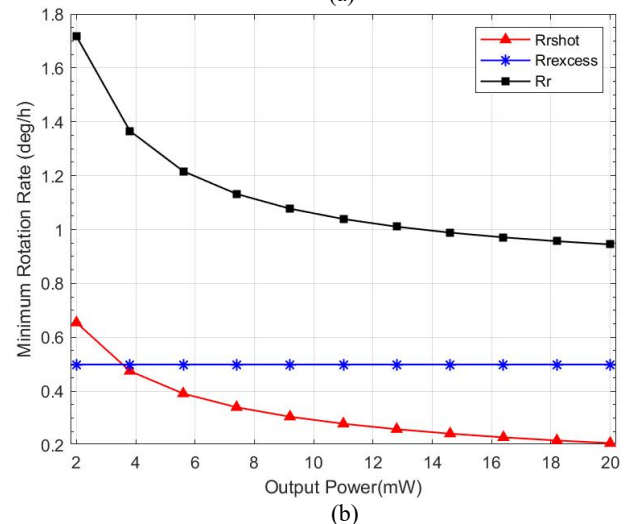
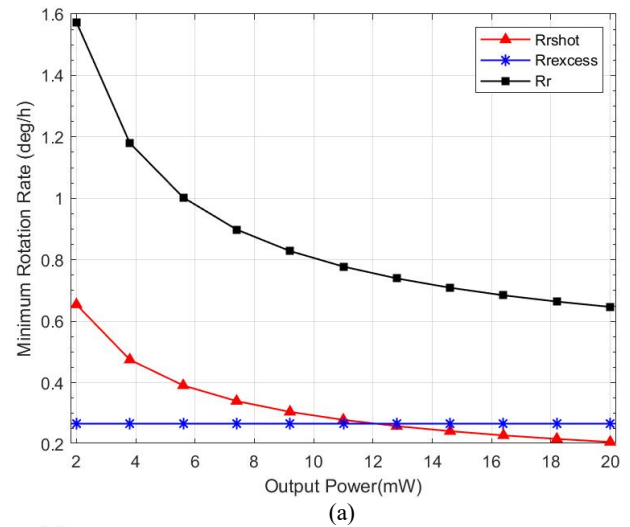


Figure V. Minimum rotation rate vs output power for a) SLD and b) SFS

The total DMRR related to SLD and SFS changes in the ranges of 1.573°/h - 0.646°/h and 1.717°/h - 0.945°/h, respectively, when the output power varies in the range of 2 mW - 20 mW.

As it is seen from Figure V, DMRR shows a falling tendency as the output powers of the optical source increase. In designs, sources with higher output power can be selected for reaching the lower minimum rotation rate, on the contrary, increasing the output power causes a large amount of energy consumption and high cost. For this reason, sources with optimum output power should be preferred by evaluating other parameters, as well.

At the point where the output power is 9.2 mW, DMRR decreases by 80 % and gets the values of 0.829°/h and 1.077°/h for both optical sources, respectively. This decreasing results in 80 % approximately, improvement in DMRR for given value of output power for both sources.

#### IV. Conclusion

In this study, the effects of photon shot and excess noises that limit the measurement sensitivity and accuracy of I-FOGs on DMRR have been analyzed for a basic configuration employing SLD and SFS. In addition, considering these noise effects, simulations with respect to the relations between DMRR and the parameters of fiber length, fiber coil diameter, photodetector bandwidth, output power, and spectral bandwidth of optical sources have also been performed.

In I-FOGs, the measurement sensitivities show an increase when the fiber length (L) and the coil diameter (D) have higher values. However, choosing convenient values for both parameters is important for optimum design by reason of the increments of dimensions and weights of the system and optical attenuation. Increasing bandwidth causes a rise in DMRR in the system. In these kinds of designs, although choosing a low-bandwidth detector enables lower rotation rate measurement, high-bandwidth devices are preferred for sampling at high rates and increasing the performance of closed-loop systems. Spectral bandwidth and DMRR are inversely proportional for the reason that the increment of spectral bandwidth reduces the effect of excess noise. The negative effects of excess noise in SFS with narrow spectral bandwidth have been also realized in other simulations performed in this study. This situation shows that spectral bandwidth has a vital role and is an important parameter in I-FOG designs. The use of optical sources with high output powers minimizes the photon shot noise and provides lower rotation rate measurement. Thus, the measurement sensitivity can be increased by using a source with high output power, but, in this manner, cost-effectiveness diminishes, and power consumption shows an increase. Therefore, it is necessary to determine the optimum values in accordance with the application.

In this study, when considering the system parameters for the optimum design as DL is 150 m<sup>2</sup>, B is 60 Hz, P is 9.2 mW and spectral bandwidths are 100 nm and 40 nm for SLD and SFS, consecutively. DMRR has been computed as 0.793°/h and 0.910°/h, respectively.

Consequently, employing optical sources with high spectral bandwidth and high output power is important for I-FOG designs to provide high accuracy and more precise measurements in navigation systems. When viewed from this aspect, this study will guide similar studies in the field and contribute to the researchers for future investigations.

#### References

- [1] Johannes Rünz, Folko Flehming, Wolfgang Rosenstiel, Michael Knoop, "Requirements and Evaluation of a Smartphone Based Dead Reckoning Pedestrian Localization for Vehicle Safety Applications", *Advanced Microsystems for Automotive Applications*, Springer, 2016.
- [2] Debeunne Cèsar, Vivet Damien, "A review of visual-LiDAR fusion based simultaneous localization and mapping", *Sensors*, Volume 20, Issue 7, 2020.
- [3] Tsunehiko Imamura, Tomohiro Matsui, Masaroni Yachi, Hideo Kumagai, "A low-cost interferometric fiber optic gyro for autonomous driving", In *Proceedings of the 32nd International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS+ 2019)*, (pp. 1685-1695), 2019.
- [4] Chris Goodall, Sarah Carmichael, Bob Scannell, "The battle between MEMS and FOGs for precision guidance", *Analog Devices Technical Article*, MS-2432, 2013.
- [5] Hyang Kyun Kim, Michel J. F. Digonnet, Gordon S. Kino, "Air-Core Photonic-Bandgap Fiber-Optic Gyroscope", *Journal of Lightwave Technology*, Volume 24, Issue 8, 2006.
- [6] Oguz Celikel, Ferhat Sametoglu, Huseyin Sozeri, "Optoelectronic design parameters of interferometric fiber optic gyroscope with LiNbO<sub>3</sub> having north finder capability and earth rotation rate measurement", *Indian Journal of Pure & Applied Physics*, Volume 48, Issue 6, 2010.
- [7] Ramón José Pérez Menéndez, "IFOG and IORG Gyros: A Study of Comparative Performance", *Gyroscopes-Principles and Applications*. IntechOpen, 2019.
- [8] López-Higuera José Miguel, *Handbook of Optical Fibre Sensing Technology*, Wiley, England, 2002.
- [9] Emirhan Sağ, Oğuzhan Coşkun, Güneş Yılmaz "Modelling, simulation and balancing of a car direction with fiber optic gyroscope and fuzzy logic algorithms", In *2019 11th International Conference on Electrical and Electronics Engineering (ELECO)*, (pp. 427-431), IEEE, 2019.
- [10] Francis. T. S. Yu, Shizhuo Yin, Paul B. Ruffin, *Fiber Optic Sensors*, CRC Press, USA, 2008.
- [11] Mario N. Armenise, Caterina Ciminelli, Francesco Dell'Olio, Vittorio M.N. Passaro, *Advances in Gyroscope Technologies*, Springer Science & Business Media, Germany, 2010.
- [12] Gerd Keiser, *Optical Fiber Communications*, Tata McGraw-Hill Education Private Limited, India, 2008.
- [13] Emmanuel Desurvire, *Erbium-doped fiber amplifiers: principle and applications*, John Wiley & Sons, Inc., Canada, 2002.
- [14] William K. Burns, Robert P. Moeller, Anthony Dandridge, "Excess noise in fiber gyroscope sources" *IEEE photonics technology letters*, Volume 2, Issue 8, 1990.

# How Does the after-COVID-19 “ABCDEF” effects model affect the development of Internet of Things and its Applications to improve Customer Experiences?

Prof. Ir Spencer Li  
Co-founder & CTO of Smart Business Consultancy Limited  
Professor of Hong Kong Adventist College  
Guest Lecturer, The University of Hong Kong  
Hong Kong, China  
spencer@smartbusiness.com.hk

**Abstract - Based on the author’s ‘after-COVID-19 “ABCDEF” effects model’<sup>1</sup> defining “architectural framework for the decision-making process,” the paper examines how human factors and emerging technologies affect organizational behaviour on implementing the digital transformation of business processes by the adoption of the Internet of Things and its applications in pandemic corona diseases. Recently, ‘COVID-19 has radically changed the global economy by accelerating the digital transformation to create New Normal customer experiences (CX). This paper summarizes “next experience” initiatives by applying six pivotal elements of the ‘after-COVID-19 “ABCDEF” effects model - Artificial Intelligence, Blockchain/Big Data, Customer Experience, Digital Transformation, Emotion, and Fintech.’<sup>1</sup>**

*Keywords - customer experience, digital transformation, internet of things, IoT, after-COVID-19 “ABCDEF” effects model*

## I. Introduction

The paper forecasts the impact and future development trend on the Internet of Things and its applications based on the author’s ‘after-COVID-19 “ABCDEF” effects model (Figure I). ‘The after-COVID-19 “ABCDEF” effects model selects and groups six contemporary emerging technologies into a framework for the decision-making process in the next normal (new normal) after COVID-19.’<sup>1</sup>

The paper examines how the Internet of Things and its applications transform business with a better customer experience. COVID-19 has radically changed the global economy by accelerating the digital transformation to create New Normal customer experiences (CX).

Internet of Things and its applications assist to achieve an excellent customer journey through designing customer journey by customer journey mapping. The tri-patriate reciprocal driving forces amongst customer care, customer experience, and customer-centricity are shaping customer behaviour with satisfaction (Figure II)<sup>2</sup>.

This paper summarizes “next experience” initiatives by examining IoT and its applications using ‘Artificial Intelligence, Blockchain/Big Data, Customer Experience, Digital Transformation, Emotion, and Fintech in the after-COVID-19 “ABCDEF” effects model.’<sup>1</sup>

The after-COVID-19 “ABCDEF” effects model categorizes its six pivotal elements as “three layers - human factors, data analytics, and emerging technologies.”<sup>1</sup> All corporate

management should consider these three layers in the “next normal” era.

### A. Customer Experience

A recent report from McKinsey emphasized optimizing customer journeys better than merely focusing on touchpoints. Customers expect to receive an excellent end-to-end customer journey with clear customer experience (CX) goals. McKinsey’s research concludes that “customer journeys are more strongly correlated with business outcomes rather than touchpoints.”<sup>2,3</sup>

The multiplier effect of each highly satisfied touchpoint decreases cumulatively the end-to-end customer journey drastically (Fig. III).

### Successful Customer Journey

‘The CX pyramid model (Fig. IV) defines clear roles of customer care vital for delivering favorable business outcomes together with customer satisfaction.’<sup>2</sup>

Data analytics can help corporates to design repeatedly “customer journey to achieve business growth and improve efficiency.”<sup>2,4</sup>

“By understanding how operational factors such as speed and first-call resolution translate into customer satisfaction, contact centers can ensure they focus their energy and resources on areas that have the most significant impact on the customer experience.”<sup>2,5</sup> If the companies focus on delivering customer journeys with solid business metrics, higher successful business cases such as higher referral rates, recurring revenues are anticipated.

“By analyzing data points associated with touchpoints, the corporates can save a lot of resources to implement new customer journey.”<sup>2,5</sup> If the companies focus on delivering customer journeys with solid business metrics, higher successful business cases such as higher referral rates, recurring revenues are anticipated.

Customer service and digital transformation are interacting to define a good customer journey. Most corporates management prefers to choose omnichannel customer journey. Daily customer behaviour monitoring can build an effective and efficient omnichannel delivering customer care in the customer journey. In the meantime, emerging technologies can build customer-centric strategies with tailor-

made solutions to deliver organizational goals. “In an omnichannel world, customer care is increasingly becoming a significant factor”<sup>2,5</sup> affecting customer satisfaction. The management is ought to build big data, captured in the customer journey, developing customer experience strategies to achieve business or corporate goals.

Customer care is always the core deliverable in the customer journey. Organization structures and functional touchpoints must be interdependent and mutually supportive.

Gartner’s research said “Eighty-one percent of customer experience (CX) leaders would compete mostly or entirely on CX. Just less than half have believed CX can help the organization to drive business outcomes. Although CX aims to deliver goods and services exceeding customer expectations, there is only 48 percent rate their CX efforts exceed management expectations and only 22 percent say that their CX efforts to exceed customers’ expectations.”<sup>2,5</sup>

“To address this challenge, Gartner unveiled the CX Pyramid, a new methodology to test organizations’ customer journeys and forge more powerful experiences that deliver higher customer loyalty and brand advocacy.”<sup>2,6</sup>

### **CX Pyramid**

The Gartner CX Pyramid (Fig. IV) lays out “a framework of from bottom to top customer experience anticipated. It can help to build an incrementally relationships between an organization’s brand and its customers based on the way CX leaders listen for, understand, act on and solve customer needs.”<sup>6,15</sup>

### **B. Internet of Things and its Applications**

Nowadays, 5G or above and the Internet of Things (IoT) are inter-operate with each other and accepted by the government, corporates, and customers. A recent Gartner survey identified the executives believing “emerging technologies creating an impact on your customer experience projects.”<sup>7,16</sup> IoT is one of the initiatives recommended for CX leaders to consider.

“By 2030, it forecasts that 125 billion devices would be connected through IoT. IoT will affect all kinds of daily lives such as B2C and B2B.”<sup>7,16</sup>

With the growing demand for IoT needs, the applications of IoT can be applied for nearly all industries. Scientists and Subject Matter Experts are working hard to design various business models of deploying IoT with specific applications.

We expect areas to be applied in the following sectors: government, NGO/NPO, MNC, large corporates, public utilities, agriculture, transportation, academics, professional firms like lawyers, SMEs, and individuals.

The critical success factors for IoT and Application development are automation, big data, connectivity, and business models.

Recent research said that due to technological breakthroughs and engineering improvements on wearable sensors, “smartwatches are being widely used in healthcare in the next five years. Data, collected by smartwatches, can be used for early diagnosis and remote patient monitoring.”<sup>8</sup>

For the finance sector, it is very convenient to apply IoT to carry out customer services and Know Your Client (KYC) procedures.

Refer to the paper Bank 2.0 - The big shift<sup>9</sup>, most financial institutes have deployed Bank 2.0 channel architecture to eliminate too many isolated banking and finance systems running together. Gradually, Bank 4.0 will come in stage to replace all banking systems by “Banking Everywhere but Never at a Bank.”<sup>10</sup> The paper predicts that IoT-based appliances can be used for implementing Bank 4.0 initiatives. For the banking sector, IoT plays a vital role in the business transformation from Bank 1.0 to Bank 2.0, and even to Bank 4.0.

On the investment side, more institutional investors, private equity, and venture capitals are keen to invest in earlier IoT projects.

In global investments, China has rapidly expanded its smart cities and IoT deployment for its large area of continents. The United States has advocated its ever-largest investment plan on infrastructures like high-speed railway, and emerging technologies like 5G, computer chipset research and development, and manufacturing.

### **IoT-based Emerging Technologies**

Due to the implementation of 5G and faster Wi-Fi, the speed and connectivity are no longer technical issues for linking IoT-based appliances. The paper has examined the trend on IoT-based emerging technologies; we conclude the following five emerging technologies enabling IoT appliances as below:

- Battery-Free Sensors
- Containers for MCU Devices
- Mesh Sensors
- Network Slicing for IoT Applications
- Small Machine Learning Sensors

#### **Battery-Free Sensors**

Battery replacements are severe problem of using IoT devices. Recently, a US-based startup called “Everactive is developing a battery-free Eversensor (Fig. V) of 20 years battery life. The sensor can accept power from combined power sources of solar, thermal, and vibration energies.”<sup>11</sup>

#### **Containers for MCU Devices**

Due to the wide adoption of IoT applications, the huge demand on “microcontroller units (MCUs) driving to design and manufacture small, low-power consumption chips supporting tens of millions of IoT appliances.”<sup>11</sup>

### **Mesh Sensors**

The wearables of the future would track the exact movement of the body rather than measuring heart rate, exertion, and sleep quality. They can be applied with new types of “garments which can collect sufficient data points which can be used by an app to determine the body’s position in 3D space.”<sup>11</sup>

### **Network Slicing for IoT Applications**

Network slicing is the technique to apply different “latency, reliability, and bandwidth for IoT devices. Network slicing is recommended to deployed on mission-critical IoT devices for long-term reliability.”<sup>11</sup>

### **Small Machine Learning Sensors**

Machine learning in sensors comprises a “growing field of hardware and software technologies that enable machine learning algorithms running on small, low-power devices like microcontrollers. The sensors can enable sensors with predictive maintenance capabilities for manufacturing equipment and super-accurate location tracking for goods on the move.”<sup>11</sup>

### **Development on the Internet of Things and its Applications**

The world is changing rapidly. There is more than 1 billion population without a bank account – categorized as unbanked population. Fintech companies are investing in developing winning APP to attract 1 billion unbanked population as customers.

Recent production shortage of computer chips let all developed, and developing countries invest in computer chip research and development, and align with countries producing chipsets to deliver steady supplies of chipsets.

Resulted from the laws of supply and demand, and technology trends, IoT and applications are critical success factors for countries to dominant in the coming few decades.

Anticipated developments on IoT and applications are

- IoT, connecting an intelligent network of smart ATMs, is the bridge between banks and their customers. IoT is one of the big data sources for customer profiling, customer services, and content management in financial services;
- Financial giants like BlackRock/PayPal are investing 10x faster pace on IoT and Future Technologies and unbanked population (fund size more than US\$130M for each investment);
- Unlock the value of IoT-enabled appliances and applications to help customer service experiences;
- Five emerging technologies enabling winning “capabilities for IoT devices and applications, from battery-free sensors to 5G network slicing”<sup>11</sup> in tiny devices, play vital dominance on IoT development

Although the future of IoT is promising, there are “international ethical concerns in Artificial Intelligence systems which are running on big data and, or IoT-based sensors.”<sup>12</sup>

The manufacturers should be aware of ethical issues violating National legislation, international conventions, particularly in the United Nations level (Fig. VI).

The paper studies the in-depth implementation of IoT to enhance automation like RPA and Fintech in advanced engineering, technology, and applications.

The paper has identified the most appropriate applications of IoT to utilize the after-COVID-19 “ABCDEF” Effects Model rationales for New Normal businesses. The matching metrics (Table I) point out the implementation areas for “ABCDEF.”

### **C. The after-COVID-19 “ABCDEF” effects models**

As mentioned before, the after-COVID-19 “ABCDEF” effects model with six elements is shown in Fig VII.

Regardless of digital transformation, ‘the important psychological “human factor” is always the first layer of the model.’<sup>1</sup> “Customer Experience” studies how to achieve a better customer experience by designing a better customer journey. Social media is an effective way of influencing and motivating Generation Z and younger generations.

“Decision-makers, such as management, scientists, and customer experience officers (CXO), must take serious consideration of the human factors, international rules, and regulations, such as the European Union’s General Data Protection Regulation (GDPR) (European Union). The GDPR is the strictest privacy and security law in the world and it imposes huge fines on those who violate its privacy and security standards, with fines amounting to tens of millions of euros.”<sup>1</sup>

“Decision-makers, such as management, scientists, and customer experience officers (CXO), must pay attention to the human factors of using data analysis, international rules and regulations, such as the European Union’s General Data Protection Regulation (GDPR) (European Union). The GDPR is the strictest privacy and security law in the world and it imposes huge fines on those who violate its privacy and security standards, with fines amounting to tens of millions of euros.”<sup>1</sup>

“The United Nations recommends the organizations migrating to cyberspace and remote participation in social, educational and economic activities which reduce the psychosocial impact of social distancing. Big data is increasingly being deployed in crisis management and predictive learning, allowing real-time data-based decision-making at a faster and more efficient response. Similarly, the world has witnessed a shift to electronic commerce over physical retail and service provision.”<sup>1</sup>

“Artificial intelligence and Big Data have been deployed on virus research, vaccine development, and data analysis for public health matters”<sup>1,14</sup>.

IoT-Based Alternative Data drives businesses to transform and adapt rapidly. IoT-Based sensors can provide emerging technologies to capture valuable data to construct special purposes big data. CX leaders should design a better end-to-end customer journey rather than focusing on a particular touchpoint. IoT and applications can add value in the end-to-

end customer journey as customers are acquainted with accepting IoT appliances as a daily life necessity.

### New Normal and Next Experience

Pandemic COVID-19 has changed the world rapidly. For hygiene and human security issues, social distancing is the norm. There is a great demand on IoT-based appliances which can provide convenient and secure services to human.

AI brings a drastic change in digital transformation for applying IoT and Applications. The new ways of O2O services can be implemented by Robotic Process Automation (RPA) and connected IoT appliances and wearables. Smart contracts generate legal contractual terms to eliminate unnecessary contractual and business exceptions in a trusted environment.

“The flexibility and agility of the Digital Service Factory make it invaluable for Japanese industrial companies’ digital transformations. By testing and proving the value of Industrial IoT and service-oriented concepts quickly, companies can get new revenue-generating digital services to market so much faster.”<sup>13,14</sup>

Five emerging technologies like battery-free sensors, containers for MCU devices, mesh sensors, network slicing for IoT applications, and small machine learning sensors will dominate the commercial and consumer markets. They drive investors to shift their investment from traditional industries into Industrial IoT to manufacture these technologies to lead the market very soon.

Humans will enjoy different kinds of living and customer journeys in the New Normal age.

## II. Figures and Tables

### A. Figures

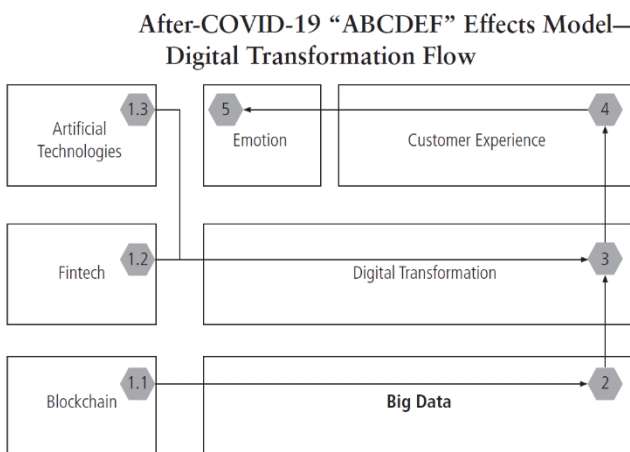


Figure I. “after-COVID-19 “ABCDEF” Effects Model – Digital Transformation Flow Successful customer journey”<sup>1</sup>

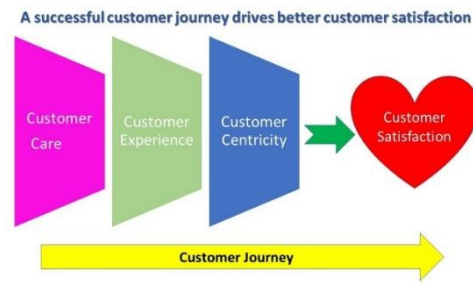


Figure II. “Successful customer journey”<sup>2</sup>



Figure III. “End-to-End customer journey and touchpoint”<sup>2</sup>



Figure IV. “The CX Pyramid, Gartner, July 2018”<sup>2</sup>



Figure V. Eversensor

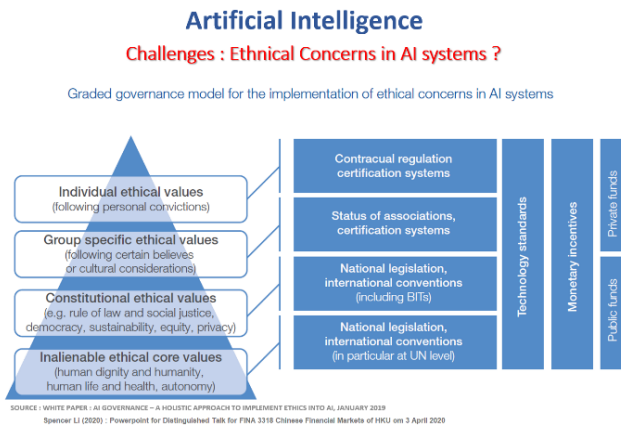


Figure VI. Challenges: Ethical Concerns in AI systems, 3 April 2020, HKU Talk, Spencer Li

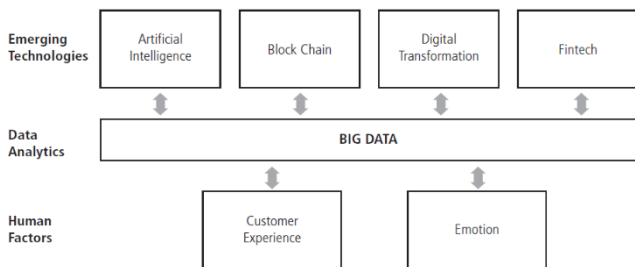


Figure VII. “after-COVID-19 “ABCDEF” Effects Model - System Architecture”<sup>1</sup>

**B. Tables**

Table I. Metrics of after-COVID-19 “ABCDEF” Effects Model affecting IoT and Applications development

“ABCDEF” Effects	IoT and its Applications
Artificial Intelligence	AI brings a drastic change in digital transformation for applying IoT and Applications. The new ways of O2O services can be implemented by Robotic Process Automation (RPA) and related IoT appliances and wearables. Smart contracts generate legal contractual terms to eliminate unnecessary contractual and business exceptions in a trusted environment.
Blockchain & Big Data	With the open APIs of common data on the blockchain and big data from different stakeholders like commercial corporates, government, and NGO, AI and ML tools can foster the development of hybrid blockchain, which deliver sophisticated services for multi-disciplines and cross-industries at a quicker pace.  IoT-Based Alternative Data drives businesses to transform and adapt rapidly.
Customer Experience	IoT and its Applications can deliver good customer care and customer-centricity services.

	Customer care is the core element driving a better customer experience. Customer-centricity focuses on the understanding of customers’ needs to derive customized strategies and solutions.  Many corporates organizations have developed omnichannel of integrating different channels using IoT technologies for sales and marketing, billing, and call centres
Digital Transformation	“The flexibility and agility of the Digital Service Factory make it invaluable for Japanese industrial companies’ digital transformations. By testing and proving the value of Industrial IoT and service-oriented concepts quickly, companies can get new revenue-generating digital services to market so much faster.” <sup>13</sup>  IoT and Applications are becoming driving forces on digital transformation.
Emotion	Low-cost wearable sensors, such as wristbands, are not uncommon used by medical and health professionals to capture blood pressure, heart rate, and temperature to feed into programmable systems to define an individual’s emotional states.  For those business transactions need peculiar focused attention and decision, emotion detection technologies have been applied to increase their attention and performance while to improve healthier productive workplace.
FinTech	An unprecedented most significant number of new business models are driven by IoT, which can be used for implementing Bank 4.0 initiatives to let customers enjoy banking services everywhere but never at a bank.

**III. Conclusion**

Development on the Internet of Things and its applications are fast. Globally, all developed and under-developed countries are building their 5G infrastructure to upgrade themselves as Smart Cities.

IoT-based sensors are easier to capture all data for data analytics purposes. Resulted of the pandemic COVID-19, new preventive measures like social distancing are taken. The author has summarized the trend of human and business behaviors in the “after-COVID-19 “ABCDEF” effects model.”<sup>1</sup> The paper anticipates IoT and applications play a vital role in the evolution of customer journey by deploying customer journey mapping. The tri-patriate reciprocal driving forces amongst customer care, customer experience, and customer-centricity are shaping customer behaviour with satisfaction.

From the macro side, huge institutional investments and government funds are in the market to seek a good investment of earlier emerging technologies. Great demand on skillful human capital and technology transfer like IP, patents are the hot topics for academics and top corporate management to consider and study.

We believe that IoT and its applications can add value to six pivotal elements of “the After-COVID-19 “ABCDEF” effects model - AI, Blockchain, Big Data, Customer Experience, Digital Transformation, Emotion, and Fintech.”<sup>1</sup> The world is so fantastic. By 2030, humans will be connected and served

by 125 billion IoT devices with new emerging technologies deployed at an unprecedentedly fast pace.

### References

- [1] Spencer Li. 2021. "How does COVID-19 Speed the digital transformation of Business Processes and Customer Experience?", SPECIAL ISSUE IN FINTECH OF "REVIEW OF BUSINESS" (St. John's New York), 41(1), 1-14, 2021. <https://www.stjohns.edu/sites/default/files/uploads/Review-of-Business-41%281%29-Jan-2021.pdf>
- [2] Spencer Li. 2021. "How Does Digital Transformation Improve Customer Experience?", The Palgrave Handbook of Fintech and Blockchain, 487, Jun 2021. DOI: 10.1007/978-3-030-66433-6\_21.
- [3] Wray, Sarah. 2016. "Optimize Journeys Not Touchpoints – Here's Why and How." McKinsey & Company. <https://inform.tmforum.org/customercentricity/2016/09/mckinseyoptimize-journeys-ottouchpoints-heres/>. Accessed September 2016.
- [4] Glagowski, Elizabeth. "The Race Is On: Rethinking Your Digital Strategy." ttec.com. <https://www.ttec.com/articles/digital-customer-experience-strategysix-key-areas-focus-your-efforts>.
- [5] Lotz, Stephanie, Julian Raabe, and Stefan Roggenhofer. 2018. "The Role of Customer Care in a Customer Experience Transformation." McKinsey & Company. <https://assets-prod.mckinsey.com/~media/McKinsey/Business%20Functions/Operations/Our%20Insights/The%20role%20of%20customer%20care%20in%20a%20customer%20experience%20transformation/The-role-of-customer-care-in-a-customer-experience-transformation-vf.ashx>. ok, SYZ Publishers, Turkey, 2005.
- [6] Kelly Blum. 2018. "Gartner Says Customer Experience Pyramid Drives Loyalty." Gartner Inc. <https://www.businesswire.com/news/home/2018073005056/en/Gartner-Customer-Experience-Pyramid-Drives-Loyalty-Satisfaction>.
- [7] Gartner, Inc June 2020. "How to Leverage the Top 5 CX Trends in 2020." Gartner Inc. <https://www.gartner.com/en/conferences/apac/customer-experience-australia/gartner-insights/gc-rn-top-cx-trends>
- [8] Dominic Hasler. 2021. "Looking after the health of your ATM fleet in a futuristic way." [www.fintechfutures.com](http://www.fintechfutures.com), 2021. <https://www.fintechfutures.com/2021/03/looking-after-the-health-of-your-atm-fleet-in-a-futuristic-way/>
- [9] Peter Müya H. 2012. "Bank 2.0 - The big shift." <https://www.slideshare.net/themuyas/bank-20-the-big-shift>
- [10] Brett King. 2019. "Bank 4.0: Banking Everywhere but Never at a Bank", 2019
- [11] Dylan Martin. 2021. "5 Emerging IoT Technologies You Need To Know In 2021." [www.crn.com](http://www.crn.com), 2021. <https://www.crn.com/news/internet-of-things/5-emerging-iot-technologies-you-need-to-know-in-2021>
- [12] World Economic Forum. 2019. "White paper : AI Governance – A holistic approach to implement ethics into AI." January 2019
- [13] Kyoko Tamur, Raghu Gullapalli. 2019. "The Secret to Maximizing the Industrial IOT." Accenture, 2019. [https://www.accenture.com/\\_acnmedia/pdf-108/accenture-apac-insight-dsf-tamura-final-lowres.pdf](https://www.accenture.com/_acnmedia/pdf-108/accenture-apac-insight-dsf-tamura-final-lowres.pdf)
- [14] United Nations Industrial Development Organization (UNIDO). 2020. "COVID-19 Implications and Response— Digital Transformation and Industrial Recovery." Vienna, Austria: UNIDO. <https://tii.unido.org/news/covid-19-digital-transformation-industrial-recovery>
- [15] Businesswire. 2018. "Gartner Says Customer Experience Pyramid Drives Loyalty, Satisfaction and Advocacy" [businesswire.com](http://businesswire.com), 2018. <https://www.businesswire.com/news/home/20180730005056/en/Gartner-Customer-Experience-Pyramid-Drives-Loyalty-Satisfaction>.
- [16] Capgemini. 2020. "Capgemini Customer Experience" [capgemini.com](http://capgemini.com), 2020. <https://www.capgemini.com/service/digital-services/customer-experience/>



# Drawing as a Scientific Method. The School of Agricultural Engineers in Madrid: a case study

Jara Muñoz-Hernández  
School of Architecture (ETSAM)  
Polytechnic University of Madrid (UPM)  
Madrid, Spain  
[jara.munoz@upm.es](mailto:jara.munoz@upm.es)  
<https://orcid.org/0000-0003-2530-2892>

**Abstract**—In the current debate on the application of new media in the documentation of architecture, the need to preserve the values of the architectural drawing tradition stands out. This paper proposes its core role in the concept of *graphic reconstitution* as a method of integration of new and traditional media for the advancement in knowledge and dissemination of architecture. A research - applying this method - on the School of Agronomists in Madrid is used to exemplify this theory.

**Keywords**—architectural drawing, graphic reconstitution, School of Agricultural Engineers, Madrid's Ciudad Universitaria

## I. Introduction

For Architecture, drawing is, inevitably, the instrument of thought, concretion and communication. From the first moment of ideation of an architectural element to the details resolution on the construction site, the entire project process goes through being drawn [1]. This paper aims to analyze how drawing, in addition to being the language of architecture – and many other technical disciplines –, is also an absolutely effective scientific instrument when it comes to researching lost or modified architectural heritage and even the architecture that was devised, but never built.

This graphic methodology has been applied in my PhD thesis, which deals with the birth and development of Madrid's School of Agricultural Engineers, the first university complex to be built on what is today the main campus of the capital, the Ciudad Universitaria. Based on everything analysed in the thesis, the aim here is to synthetically present the method, take the School of Agronomists as a case study and show the results obtained, in such a way that it can also be applied to other areas of study.

This research is not an isolated work, but is included within a broader research framework related to the drawing of the city and architecture, from which various scientific articles [2] and several PhD theses have been carried out in the old research group of the Technical School of Architecture of Madrid, *Drawing and Documentation of Architecture and City*. All of them are studies that start from a work begun a long time ago and whose first verifiable result is the book *La forma de la villa de Madrid* [3].

## II. A case study: The School of Agricultural Engineers

Nowadays, the northwest corner of Madrid is unfailingly linked to the Ciudad Universitaria. The social, political, urban and symbolic value of the campus has ended up by imposing its name over this whole area of the city, relegating to oblivion – or reserving them for specific places – the names bounded to the origin of this space: Florida and Moncloa.

In 2019 the School of Agronomical Engineers celebrated its 150 anniversary. It was the first institution established on these territories, coinciding in time with the end of the Royal

Estate of La Florida and La Moncloa. Its status as a Crown property had indeed preserved the estate from urban growth and it would also be so in the following decades, once the State decided to transfer it to the agronomists, thus starting its educational and research trajectory. Today the School occupies a small part of the campus, but in 1869 the entire La Florida estate had been surrendered to the institution. The grounds would be later occupied by various charity, sanitary and recreational centers. La Moncloa was hence configured as a natural park facing Casa de Campo, performing a transition between the city – which had reached its limits with the construction of Argüelles neighborhood – and the mount of El Pardo. The fusion of the cultivated land and the public garden, together with the distinctive topography of the landscape, gave this area a picturesque character that experienced great success among the people of Madrid.

La Moncloa as a Royal Site until 1869 and as a University campus since 1927 has already been studied. However, that time in between those two dates has only been partially researched in several papers. Consequently, its architectural and urban aspects had not been approached in a global way. It is yet to be understood how the territory and pre-existing constructions where occupied after the transfer of the State, how it was developed during sixty years until the creation of the Ciudad Universitaria and how this project coexisted with the School of Agronomists and other institutions until the Spanish Civil War (1936-1939). Unfortunately, the excessive and uncontrolled growth since the last sixties, almost completely erased the traces of the past.

This paper aims to approach all these questions through drawing, which is understood as a source of information, an instrument of analysis and a means of expressing results. In this way, graphic narration becomes an essential and enriching complement to the written discourse, showing the research area at various key moments of its development. The result is a sequence of drawings of one same place at different moments in time that allows recovering – in a virtual way – the lost memory of this place, as well as contributing to the knowledge of the urban form of Madrid.

## III. Method

In order to achieve the objectives described above, systematic work in archives and libraries has been carried out. However, the main methodological instrument of this research is drawing. This work can be framed within lines of research that have shown that drawing is an effective tool in the analysis of the urban form and in the transmission of results and conclusions.

In another sense, drawings are also a source of information. Thus, our work is nourished by the documents that architecture and the city generate and that are created during their development. The set of all these drawings is what has been called the *graphic life of buildings* [4].

Most of the graphic documentary sources examined in this research come from libraries and public archives. We will not stop to analyze the origin of the sources – [5] can be consulted for more information – but it is convenient to refer to the most important ones. Obviously, the information that each file can provide is closely linked to changes in ownership of the scope of study. As La Florida was a property of the Crown, the General Archive of the Palace contains written documentation and plans from that historical period. Later, once these lands are transferred, the information is kept in other files. The General Administration Archive, an archive that preserves the documentation of the different Spanish ministries since the second half of the 19th and 20th centuries, contains information on the project of the School of Agronomists and a large part of the multitude of buildings that were built in the surroundings of the Palacete de La Moncloa. Since the 1920s, when the Construction Board of the Ciudad Universitaria was constituted, and especially after the Civil War, the General Archive of the Complutense University of Madrid is also an essential place of consultation.

The systematic collection of information is essential as a starting point, both for written work and for the production of graphic documentation. However, this compilation does not end exactly at a certain point, but research and work feed one another. It is clear in the drawing of plans, for example, that, with a good collection of old graphic documentation, writings, etc., it is possible to begin to draw, but it will always be necessary to fill in the gaps that will force a new search for documentation while those planes are being generated. It is this back-and-forth process that builds the graphical base of which we spoke before, and that must be understood not as a closed and finished product, but rather as a work in which information can continue to be poured into a constant expansion and growth process.

#### IV. Drawing as a scientific instrument

Up to now, drawing has been discussed from a documentary source point of view. However, in addition to its value as a container of information, it can also be considered as a tool for thought and analysis, as a project towards the past – of what existed or what could have existed – or as a method to illustrate and reflect the results of an investigation. All these meanings will be taken by the accompanying drawings, which *are* the research we are talking about.

It is worth stopping at the use of drawing in a scientific and rigorous way, analyzing the original documentation, whether graphic or written, and synthesizing it in a graphical, georeferenced database, which can give rise to complete planimetries, 3D models and images, in what we understand as a *graphic reconstitution* of the case study: “... the term reconstitution would be reserved for those drawings that attempt to reflect one or more states of the building that no longer exist or that never existed, but that could be part of its biography. Note that the important difference is that in the second case, given the almost always incomplete data, it is usually necessary to introduce a certain dose of interpretation we would like to assimilate to a certain idea of a project”[4].

On this same basis it is possible to take a trip to the past, in which drawing is an essential means of shaping that *graphic life* of buildings. In this process of graphical reconstitution, the most important thing is to minimize the data subject to speculative interpretation. Obviously, this uncertainty is accentuated the further one is on the timeline.

The first step for this graphic reconstitution is to set those elements which remain, that is, that existed then and that still exist today. In order to do so, they are identified on the current urban parcel, which we consider the most reliable georeferenced cartographic source. These elements are the only certainties that are available and that serve as a reference for graphic reconstitution [3]. These persistences can be buildings or constructions, that is, they have a clear material dimension, or be elements related to the orography, water courses, roads or property limits – that can manifest themselves in a variable way in different historical moments, either through built boundaries, fences, street names... –, that is, they have an immaterial character [6].

The second step would be the location, based on the reference marked by the persistence, of missing or transformed elements of which there is measurable documentation. In the case of La Moncloa, which for a long time was a sparsely built area, it happens that some buildings are built on top of others, or some parts are used for their development, which also makes it possible to take the current building as a reference to locate the old one (Figure I).

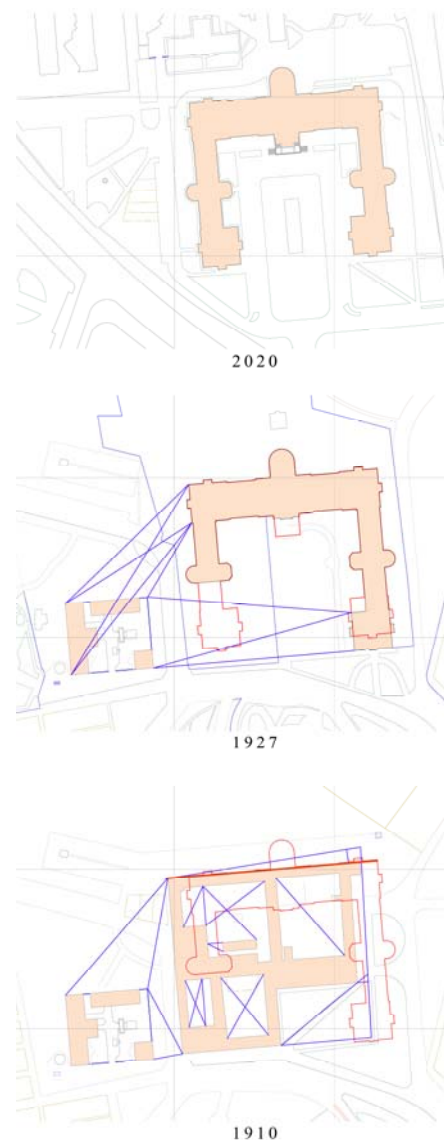


Figure I. Process of graphic reconstitution of the location and layout of the Porcelain Factory (bottom), taking as a reference the current building of the School (top), its layout in 1927 and the situation with respect to this of the Machine Building (center).

For example, in the case of the teaching building of the School of Agronomists, we have nowadays the original building partially enlarged and modified, as can be seen in the upper drawing of Figure I. From it, we can draw the layout of the original building, of which we keep plans. And, once the original building is located, it has been possible to locate one of the auxiliary buildings that existed in the 20th century and that were later demolished. This is what the middle drawing represents, where the footprint of the current building that has served as a guide is kept in red line. Finally, and following the same graphic codes, we have been able to locate the building prior to the current one of the School of Agronomists, which had been the Crown Porcelain Factory [7]. This has been possible thanks, on the one hand, to the knowledge of the situation of the secondary buildings, which coexisted in time with this construction. And on the other, to information obtained from historical photography: the Porcelain Factory and the new School of Agronomists coincided on their north façade. It was relatively easy to obtain the floor plan of the Porcelain Factory building, since we have some historical plans with measures. However, it was not so easy to place it in a space with very few built references. This is an example of how the drawing, rigorously built, allows us to obtain this information. In the same way, the plans of the complete area that will be shown later have been produced.

Naturally, matching historical cartography to one another or to the current city maps is a useless task, since the different projection systems and the distortions and inaccuracies of the survey methods at all times prevent it from being transferred directly onto the current parcel map. This problem is even more accentuated in sparsely built areas, where it was more difficult to establish benchmarks from which to measure. For this reason, in this case, references have been sought with the present, or with the more controlled past, “building by building”, in order to later be able to draw with precision those for which there was data. These data have basically been obtained from individual projects or, also, of the measurements of the polygons of some buildings that appear in the field sheets at a scale of 1: 500 that were drawn for the subsequent elaboration of the urban plot of the Statistical Board (1860-1870), which are preserved in the National Geographic Institute.

The third step of the process consists of defining those elements whose precise shape or location could not be determined from dimensional documentation, and from which there are currently no references that can be taken. It is then that one enters the field of interpretation and speculation, in a process very similar to that of an architecture project, which does not mean that there are no reference elements on which to rely, in order to establish sensible hypotheses: cartography history, written documents, press, photographs, paintings and engravings ...

Once this graphic reconstitution has been carried out and reliable drawings have been obtained, one can experiment with the tools at our disposal – the most common, newer ones, or in combination – to decide how to display these drawings.

The *comparative method* is one of the most powerful tools for analyzing and understanding the urban form of cities or, in this case, of the same city at different moments in time. By referring all the drawings to the same cartographic base, with a common scale and the same graphic variables, we can establish a “graphic parallel”: “We would thus arrive at the concept of graphic parallel, which we could enunciate as the

result of presenting, at the same time and according to the same criteria, a series of different architectural facts in order to classify or compare them. Thus, the graphic parallel becomes a preliminary step in any systematic study with a scientific vocation in which more than one architectural object is involved and which has more or less form as a direct reference” [8].

In this case study, the application of the graphical parallel has consisted in comparing the same spatial area at different moments in time. This area has been drawn on two different scales, which make it necessary to approach the drawing in a different way.

#### A. Spatial scope of the drawings and temporal sequence

From this case study, three packages of plans of the same historical sequence have been made, each one with a different spatial framing (Figure II): an urban one, at a scale of 1: 7,500, in which the territory of La Moncloa occupied by the agronomists is represented, and its relationship with the city and with structural elements of the landscape, such as the Manzanares River or the Casa de Campo. The second and third frames, both at 1: 2,000 scale, reproduce the surroundings of the School of Agronomists and the whole of the Model Farm that was developed in the vicinity of the Palacete de La Moncloa. This second close frame, due to historical circumstances, will disappear in the time scenes after the Civil War. Architecture is already represented on this scale and includes the graphic reconstitution of all those buildings for which a minimum of information has been found to establish hypotheses.



Figure II. Spatial scope of the research over Madrid’s satellite view.

Temporarily, key moments in the development of the urban form of the study area have been chosen, in which its image has been “frozen” in order to study it:

-1870, as the moment of the transfer of the property and preliminary state of the place.

-1890, a state prior to the end of the century and the appearance of institutions other than the School, the year in which the Alfonso XII Agricultural Institute was founded, the

first refurbishment works were undertaken and the expansion process started.

-1910, as a state immediately prior to the new building of the School and the date in which the development of the institution is already remarkable. This section has been drawn from the population plan of the same year.

-1927, the last moment of the School before the construction of the campus (Figure III).

-1936, date marked by the beginning of the Civil War.

-1955, a state in which the first reconstruction projects of the School building had already been undertaken.

-2020, as a current image as a result of the processes seen throughout the work.

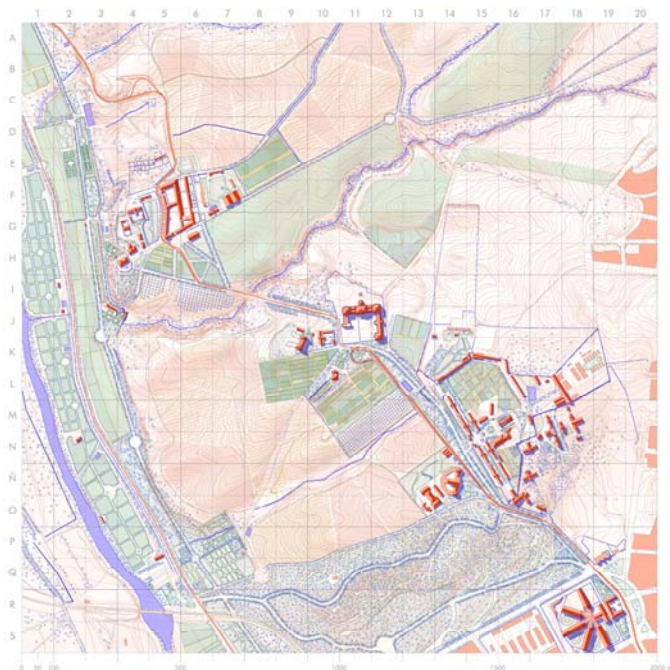


Figure III. Drawing of the research area in 1927.

## B. The drawing of the territory

On the distant scale, digital terrain models have been made to allow us to understand what the topography of Florida was like before it was modified (Figure IV). By means of these models we can establish quite reliable comparative studies of the orography before and after the construction of the Ciudad Universitaria. The current model of the land (DTM) is now available to the public; the model of the year 27 has had to be built from the contour lines of the parcel sheets of Madrid from 1929. The topographic operations prior to this date were quite minor – circumscribed to the construction of new buildings –, so this model is also valid for earlier dates. A model of the terrain from 1936 has also been developed, which makes it possible to understand which earth movements were caused by the Ciudad Universitaria project, and which have a later origin [9].

The colored images of the terrain models can be used as a basis for planimetry, as well as a source to obtain three-dimensional models. In addition to conventional three-dimensional models, we can create “fake” models from depth diagrams obtained from DTM images. In this way we obtain an apparent 3D model that is much lighter to manage.

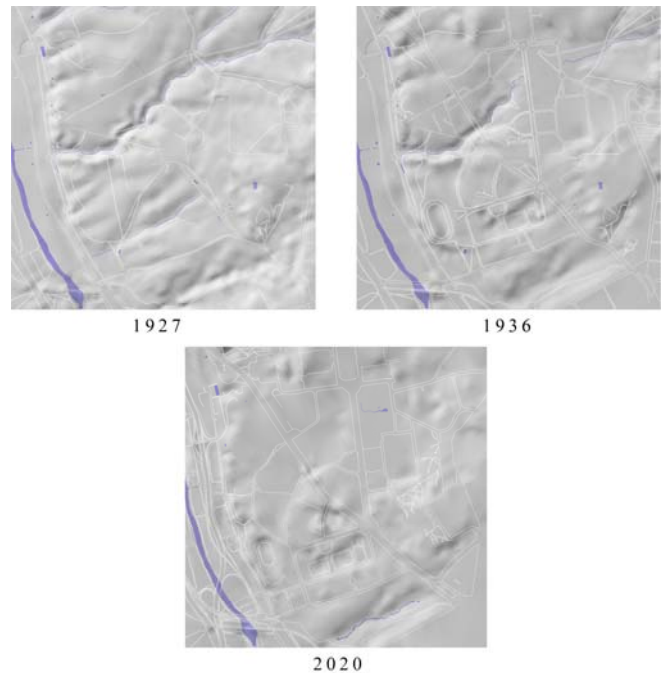


Figure IV. Topography of La Moncloa in 1927, 1936 and 2020.

In this general scale, some planimetries have been produced in each of the temporal cuts mixing the 2 and 3 dimensions. A more scientific drawing, geographically referenced, is then edited with textures and patterns and mixed with layers of 3D elements that make it possible to incorporate lights and shadows. With these plans, the aim is to achieve plans of the same scale, drawing and editing criteria, which allow a real comparison between historical moments.

## C. Property delimitations

Another function of drawing that has been basic throughout history is to establish the limits of properties. In our example, La Moncloa goes from being a property owned by the Crown to belonging to the State. Its outer limits were varied and are carefully described in the thesis [5].

Internally, the delimitation is simpler since there are no different owners. However, over the years, the State gives land to various institutions, in addition to the agronomists, which is also interesting to reflect in the drawings. This issue will be particularly relevant with the arrival of the Ciudad Universitaria, when the scale of the project is such that it invades previously occupied land, and also after the Civil War, at which time the state of ruin of much of La Moncloa and the growing needs of the Campus forces to rethink the property or usufruct limits of the territory [10].

In our case study the property issue was quite clear, but it is easy to find examples where public and private properties are intertwined in a complex framework (the historic nuclei of cities, without going any further). This topic is a good example of what the economy of drawing entails: a single graphic document allows us to understand at a glance the division that owns the territory, a topic that in writing is extremely complex and stormy, and requires the writing of a long number of pages.

## D. The drawing of architecture

On a smaller scale, the same process has been followed as for the drawing of the city or the territory. In this case, as some of the buildings that are currently preserved have been accessible, some partial surveys have been carried out, in

order to be able to compare them with the original project plans or subsequent renovations. Data collection is essential in the production of close-scale documentation. Through the first sketches, the order of the buildings, the modulation and the existence of elements that are repeated are observed.

From these plans, 3D surveys have been carried out, either modeling with a high degree of detail, to obtain descriptive drawings, or performing a simple extrusion and mapping afterwards the original plans on it, for models with a more analytical purpose. For the modeling of the School of Agronomists, the order module – including column, cornice, frieze and brick panel – has been built and the complete building reconstructed from it.

The interest of modeling goes beyond obtaining three-dimensional images of the building that no longer exists today: studying current and old photographs and placing the same cameras within the model, highly accurate comparisons can be made that allow reconstructing the landscape of the primitive Ciudad Universitaria [11].

These working methods should not be considered as separate elements, but as a set of combinable tools. Working with photography is a round trip, always in conjunction with the rest of graphic and written methods. The first step to be able to make a photographic comparison of a specific area or building is to know what has been written about it and study the different planimetries to be able to establish the time and place where the image was taken. It is important to check each photograph and identify them well, before continuing to work with them, as many times photographs with totally wrong descriptions are discovered. One of the consequences of the Civil War was precisely the destruction and disorder of the documentation housed in the faculties. The widespread chaos after the war caused many photographs to be quickly and poorly cataloged.

Once the photographed place has been identified, the two and three-dimensional drawing acquires an important role, since, through the plan plans, we can locate the position of the photographer quite roughly and later we can achieve even greater precision by means of the placement of cameras in three-dimensional models. Once located, current photographs are taken on the ground, looking for the studied point and making the last position adjustments that are necessary, provided that the current state of the place allows it.

Finally, after taking two photographs of the same place at different times, they can now be compared and analyzed in depth. The meticulous study of these photographs helps to specify details that do not appear in the plans or to adjust the modifications that were made during construction and were not reflected off-plan. They are also a valuable source of information to learn about the original finishes of these buildings and see how they have changed. The “skin” of buildings is one of the parts that undergoes most transformations over time, so this will be the most reliable document for its reconstruction.

Once the drawings have been constructed, one must think about how to display them to explain the research, that is, in its narrative dimension. Again, the role of drawing is very interesting because it allows us to establish narrative lines at different levels: absolutely descriptive of reality, as seen in some examples above, or partially descriptive, allowing us certain licenses that allow us to better show ideas. For example, in the case of the School of Agronomists, the

symmetry of the main facade of the building has been used to show, on one side, the original building and, on the other, the current one, so that, in a very synthetic way, it is possible to understand the transformation of the building (Figure V).



Figure V. Main elevation of the School of Agriculture engineers. On the left, façade in 1936 and on the right, current façade.

Another narrative level of great interest is that of analytical drawing. In this case, simpler drawings are usually used, with a lower degree of detail, but which nevertheless condense a large amount of information. In this paper two examples are offered: on the one hand, that of the constructive evolution of the School of Agronomists, where in a schematic way and with a color code the change of the building is shown since its initial construction (Figure VI).

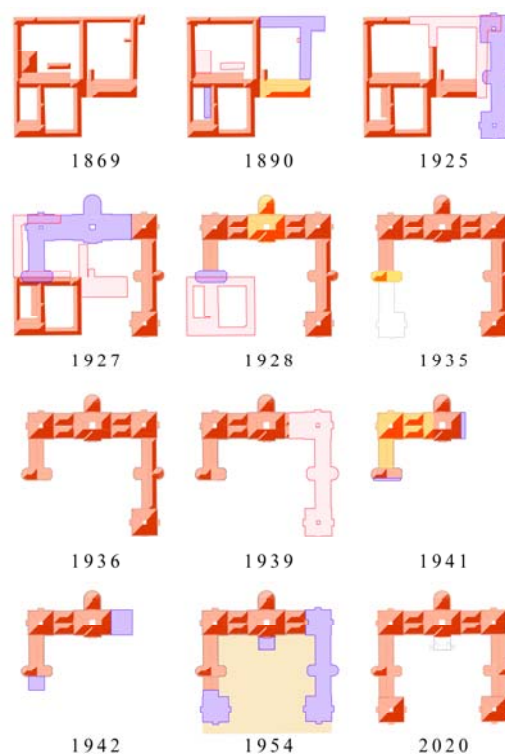


Figure VI. Evolution of the School of Agronomists from 1869 to the present.

On the other hand, thanks to the fact that all these drawings have been scientifically built, a certain precision can be guaranteed that has enabled me to quantify the surfaces built in this territory, in addition to determining which surface was destroyed during the Civil War and how much was rebuilt. The conclusion is the following: in the area dedicated to teaching, the percentage of destruction was very similar to that of the campus average – 42%, although the School of Agriculture building suffered especially – but in the Model Farm and the surroundings of the Palacete the destruction was practically total (86%) and, in addition, it was decided not to rebuild it, contrary to what happened with the teaching environment. With this decision, the School was *de facto*

incorporated into the campus as a whole and the hegemony that until that moment it had had, at least in extension, was eliminated.

## V. Conclusion

In an investigation where the drawing is an inherent part of the research, it also becomes the result and conclusions of the work. Thus, to date non-existent graphic material has been produced that represents a space in the city of Madrid at different times, in a constant journey between territorial and architectural scales.

The comparative analysis of these historical sequences allows, thanks to the unit of scale, framing and graphics, a fluid reading of the evolution of the field of study in two aspects: the spatial, which orders the elements at a specific moment, and the temporal, that orders the transformations successively (Figure VII).

As indicated above in reference to the location of already known places and constructions, the drawing is precisely the tool that has made it possible to position them in specific coordinates and graph them in the plans prepared for the thesis.

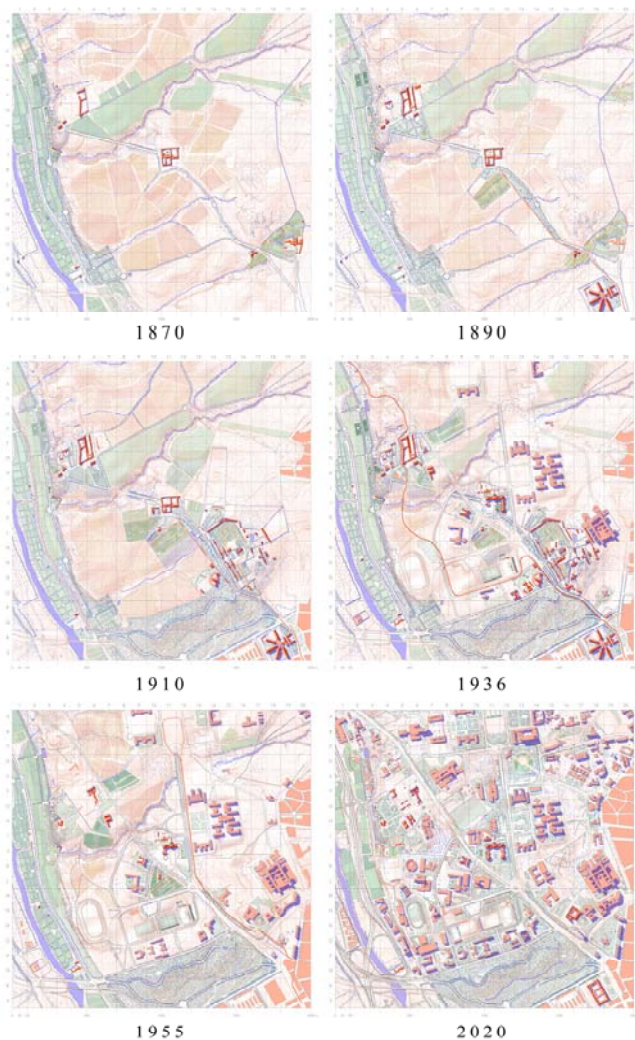


Figure VII. Drawings of the research area in each of the key dates

On the other hand, drawing has also been considered in its narrative dimension, understanding it as an essential complement to written discourse. This graphic narrative supposes, in itself, a contribution to the knowledge of the city, insofar as it provides an image of it that has not yet existed. In addition, the support of the drawings, which is not only physical on paper, but also digital, since they have been made with computerized means, generates a georeferenced database that becomes a piece of a much larger set of previous research on the city of Madrid, available for future studies. In this sense, there is also a possible further development of research in means of representation towards more applied aspects, since the abundant graphic information that has been produced here could be organized in a Geographic Information System. This path has already been started by building all the drawings on a digital common basis, with which a database could be generated in which the current campus and its previous states could be considered. This would work as a starting document for any urban development intervention, diffusion activity and even heritage recovery projects.

## References

- [1] Jorge Sainz, *El dibujo de arquitectura: teoría e historia de un lenguaje gráfico*, Reverté, Barcelona, 2005.
- [2] Jara Muñoz-Hernández, Carlos Villarreal-Colunga, “Las andanzas de la portada de Oñate tras la demolición de la casa-palacio: calle Mayor, Teatro Español, La Moncloa”, *Arqueología de la Arquitectura*, volume 17: e094, 2020. <https://doi.org/10.3989/arg.arqt.2020.003>
- [3] Javier Ortega-Vidal, Francisco José Marín-Perellón, *La forma de la villa de Madrid. Soporte gráfico para la información histórica de la ciudad*, Dirección General de Patrimonio Histórico, Madrid, 2006.
- [4] Javier Ortega-Vidal, Ángel Martínez-Díaz, María José Muñoz-de-Pablo, “El dibujo y las vidas de los edificios”, *EGA Journal*, volume 18, 2011. <http://doi.org/10.4995/ega.2011.1335>
- [5] Jara Muñoz-Hernández, *La Escuela de Ingenieros Agrónomos en La Florida-Moncloa [PhD thesis]*, Universidad Politécnica de Madrid, 2020. <https://doi.org/10.20868/UPM.thesis.65305>
- [6] Luis Sobrón-Martínez, *Al Este del Retiro [PhD thesis]*, Universidad Politécnica de Madrid, 2015.
- [7] Jara Muñoz-Hernández, “De la Fábrica de Porcelana a la Escuela de Agrónomos de Madrid”, *Revista de Humanidades*, volume 41, 2020.
- [8] María José Muñoz-de-Pablo, Ángel Martínez-Díaz, “El paralelo. Bosquejo de un método gráfico”, *EGA Journal*, volume 23, 2014. <https://doi.org/10.4995/ega.2014.2172>
- [9] José Luis González-Casas, Jara Muñoz-Hernández, “The urban and environmental Impact of Madrid’s Ciudad Universitaria: A comparison between the first Campus and the post-war Campus”, *International Journal of Sustainable Development and Planning*, volume 15, issue 6, 2020. <http://doi.org/10.18280/ijstdp.150612>
- [10] Jara Muñoz-Hernández, José-Luis González-Casas, “Traces and scars. The reconstruction of Madrid’s Ciudad Universitaria after the Spanish Civil War”, *WIT Transactions on The Built Environment*, volume 191, 2019. <http://doi.org/10.2495/STR190181>
- [11] José Luis González-Casas, Jara Muñoz-Hernández, “Drawing for heritage dissemination. The birth of Madrid’s Ciudad Universitaria”, *International Journal of Heritage Architecture*, volume 2, issue 2, 2018. <http://doi.org/10.2495/HA-V2-N2-359-371>

# Urban distribution network proposal: A case study for the 14th district of the city of Medellín.

Juan P. Vasco-Gallo  
 Ingeniering Production Student  
 Instituto Tecnológico Metropolitano  
 Medellín, Colombia  
[juanvasco241321@correo.itm.edu.co](mailto:juanvasco241321@correo.itm.edu.co)

J. Isaac Pemberthy-R.  
 Department of Quality and Production  
 Instituto Tecnológico Metropolitano  
 Medellín, Colombia  
[jorgepemberthy@itm.edu.co](mailto:jorgepemberthy@itm.edu.co)

Eduard A. Gañan-Cardenas  
 Department of Quality and Production  
 Instituto Tecnológico Metropolitano  
 Medellín, Colombia  
[eduardganan@itm.edu.co](mailto:eduardganan@itm.edu.co)

**Abstract**— Today, in the context of the world's cities, the displacement of the population to urban centers has significant consequences. This work seeks to propose a last mile distribution network, controlled by the government of a city, which seeks to reduce the impacts of this type of transport on the quality of life of the inhabitants. The main objective is to build a proposal that reduces the impacts that today, in times of pandemic, brings to the city the increase in grocery delivery service. For this purpose, an integer linear programming optimization model (MILP) was built, obtaining agile results in a first scope focused on a district in the city of Medellín, Colombia. The main theoretical contribution of this work is the development of an efficient MILP that responds to the objective under a problem in a real scenario.

**Keywords**— City logistics, Urban logistics, Operations research, Optimization.

## I. Introduction

The problem of urban distribution of goods or city logistics has been widely studied over the years and today, it is no exception since logistics is one of the most important activities in the context of business and developed cities of the XXI century. Nowadays, city logistics is becoming more noticeable, because, with the growth of e-commerce [1] and the global urbanization trends that force modern cities to offer opportunities for employment, education, culture, health, sports, among other activities such as, the development and growth of industries. This leads to the expansion of urban areas and the increase of road traffic and consequently to the increase of environmental pollution, vehicular congestion, negative social impacts, generating poor quality of life of citizens [2], in addition to an inefficient and ineffective service reducing the level of service of city logistics [3].

Today, given the current situation of confinement generated by the COVID - 19 pandemic, the demand for groceries has increased through home delivery services or through e-commerce, thus producing more trauma in the logistics of the last mile in the city [4]. As can be seen in Figure I, searches for groceries delivery services in Colombia in the last year have increased considerably because of the pandemic. Figure II and Figure III show that the state (Antioquia) and the city (Medellin) are among the most represented populations in the index of searches in the Google engine. This increase in Google® searches is a good index of the increase in the execution of services of the same type. In this way, urban logistics can have an impact on aspects such as traffic congestion in the city and the generation of vehicular pollutants.

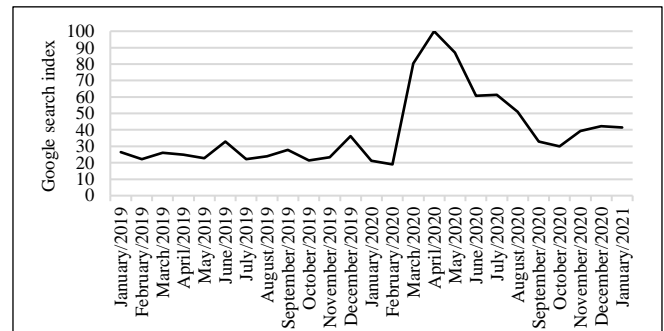


Figure I. Food domicile search trends in the Google search engine in Colombia. Source: Google Trends @ [5].

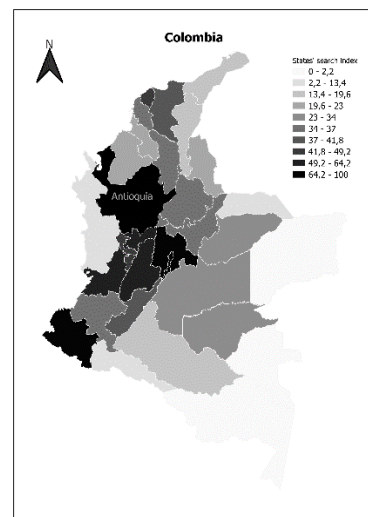


Figure II. Choropleth map of the search index of grocery home delivery services by states in Colombia. Source: Google Trends @ [5].

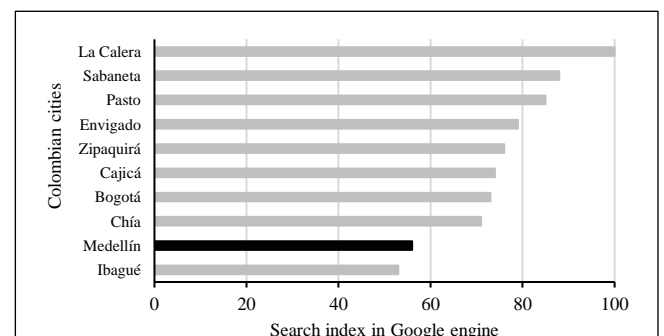


Figure III. Colombian cities with highest search rates during 2020. Source: Google Trends @ [5].

A possible strategy proposed by some authors to solve the actions of city logistics, are urban warehouses (WH) or also known as urban consolidation centers [6], whose main function is to redirect as much as possible the flow of goods and provide efficient transportation from the WH to the urban areas of the city and vice versa, through the change of long-distance cargo vehicles to short-distance vehicles [7]. As evidenced in some works, where the installation or not of urban WH is defined, applying mathematical models; this case of application of integer linear programming (MILP), to define the strategic installation of WH in the urban perimeter, where it is concluded to install satellites in the peripheries of the city [7]. Another outstanding case is based on a methodology based on descriptive survey, applying a study of a multicriteria structure for the sustainable implementation of urban WH in cities, where the results define not to install any WH in a small city in Brazil [8]. Another case study of linear programming models is applied in the characterization of the supply chain of bovine products with high production in the province of Sabana Centro (Colombia), where the model results in opening several centers in different strategic cities for the company [9].

In this work, a case study is developed in the city of Medellín, today it is ranked among the most congested cities in the world according to the INRIX index [10]. The case seeks to address a WH location problem by applying an integer linear programming model, which optimizes and responds to current needs, with respect to the distribution of basic products of the family basket. The district 14 of Medellín called El Poblado, is chosen as a case study. This district has special characteristics, such as the concentration of young population with a high socioeconomic level, which makes its population more prone to the use of technological services and therefore to the home delivery service.

## II. Materials and methods

Given the current situation, a sequence of phases has been constructed, which allow proposing an improvement solution to the problem posed, by means of the consolidation or installation of WH.

### A. Characterization and delimitation of the case of study

Medellín is the capital of the state of Antioquia and is in the Aburrá Valley, it is in the center of the state. This valley is composed of 10 cities forming the Metropolitan Area. Medellín is the most populated city in the valley, with a population of 2.5 million of the nearly 4 million inhabitants of the Valley. Medellín is located in the middle of the Valley and is divided into 16 districts comprising 249 neighborhoods as illustrated in Figure IV below.

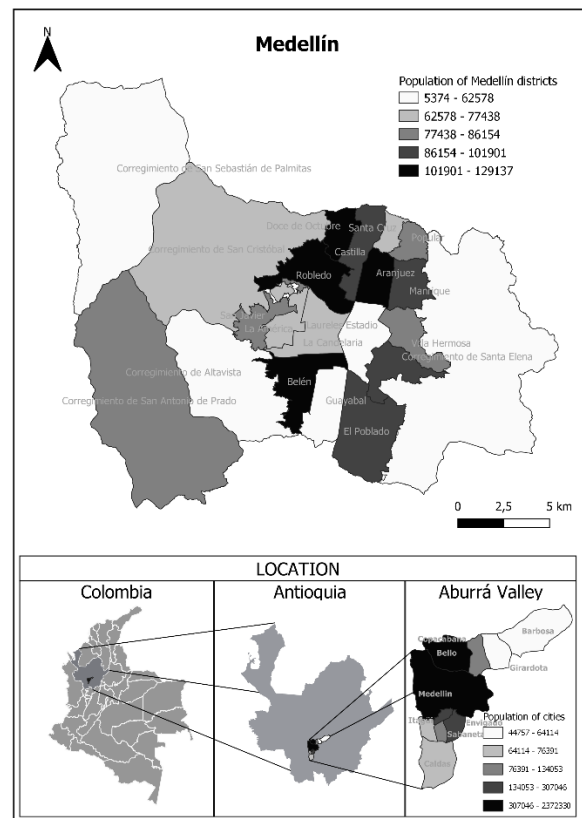


Figure IV. Medellín Districts' location and population. Source: [11],[12]

For this purpose, district 14 El Poblado is selected as the scope of the study, due to the importance of this area for the city. As it is the second district with the greatest logistic influence in the city [13]. This district has 22 neighborhoods, and its population has a decreasing growth, according to the population growth rate and demographic profile [14].

### B. Input data

To find the solution to this kind of problem, different information is required depending on the model to be applied in the development of the solution.

#### ▪ Geographical locations

Based on a guide map of the division by neighborhoods of district 14 El Poblado, an overlay of a cartesian plane is made in order to establish the coordinates (x,y) for each neighborhood under study. To do so, we define to work with their centroid points as a reference point of location. The different candidate locations for WH are defined in points that have the availability of leasing premises, or spaces for industrial facilities as WH. Given that district 14 is mainly of Urban characteristics, there are few candidate locations in this territory to locate a Warehouse. The location of the centroids of each neighborhood (red dots) together with the candidate locations for WH (blue triangles) are shown in Figure V.



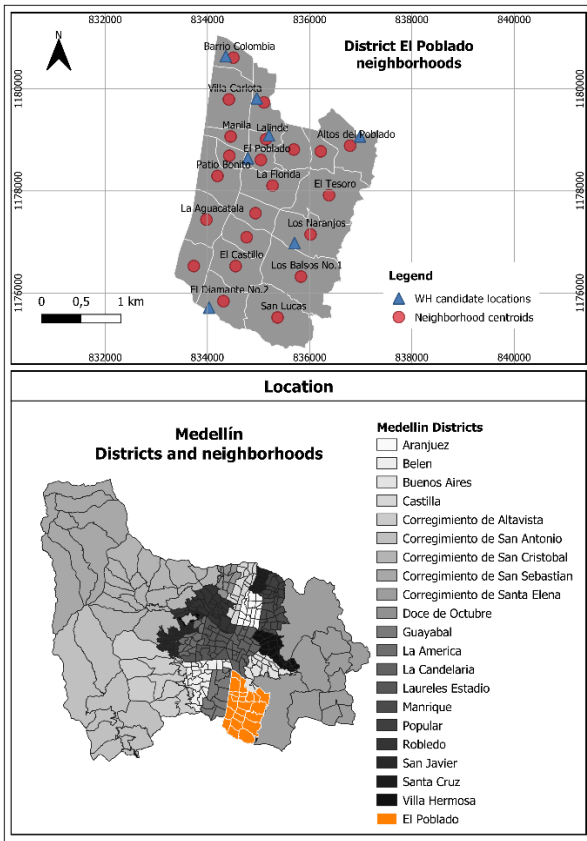


Figure V. El Poblado district's location, neighborhoods centroids and WH candidate locations.

#### Distance between locations

For this purpose, it is defined to work with Euclidean distances in the defined cartesian plane. The distances used for the case are calculated from the centroid points of each neighborhood to the different candidate locations for the location of the WH. The distances are calculated using QGIS® software [14].

#### Transportation and facility costs

The cost of shipping goods is assumed to be COP \$819.5 (\$0.22 US dollar) per kg. shipped per kilometer traveled; this cost is taken as a reference from the "Rappi® delivery" platform [16]. On the other hand, the cost for the installation is assumed as an average cost of renting premises such as warehouses or commercial places in malls or shopping centers present in the study area, it is COP \$19 million in average (\$5201,26 US dollar). This information is obtained based on the platform of the local leasing company in the city [17].

#### Population demand in each neighborhood

In this case, the population of each district will be used as potential demand, without loss of generality when using the average individual or per person consumption. This fact is used to represent of consumption when calculating the representative transportation cost for each neighborhood in the district of study. The population per neighborhood is available in [13].

### C. Mathematical statement of the problem

The model objective is to find a proposal for the opening of urban WH in El Poblado district, seeking to minimize the costs of opening and transporting basic necessities in the last mile to the various neighborhoods, and, in compliance with the assignment of each neighborhood only to a proposed WH located at a distance of no more than 2 km. For the formulation of the problem, two binary decision variables were used,  $x_j$  that establishes whether a WH is proposed to open in candidate location  $j$ . And the variable  $y_{ij}$  that establishes whether the attention of neighborhood  $i$  is assigned to the proposed WH in location  $j$ . Table I defines the terms used for the linear modeling of the problem.

Table I. Definition of mathematical model terms.

Sets	
$N$	Set of neighborhoods available in El Poblado district. $N = \{1,2,3, \dots, n\}$
$U$	Set of candidate locations for the location of urban WH. $U = \{1,2,3, \dots, u\}$
Indexes	
$i$	Identifies each of the neighborhoods in the district. Where $i \in N$ .
$j$	Identifies candidate locations throughout the district. Where $j \in U$ .
Parameters	
$n$	Number of available neighborhoods in district 14, El Poblado.
$u$	Number of candidate locations for the location of urban WH in district 14, El Poblado.
$C_{ij}$	Estimated cost of shipping one kilogram per kilometer in the city of Medellín from candidate location $j$ to neighborhood $i$ .
$P_i$	Population of neighborhood $i$
$d_{ij}$	Euclidean distance in kilometers from the centroid location of neighborhood $i$ to candidate location $j$ .
$CI_j$	Cost of installing or opening WH in candidate location $j$ .
$a_{ij}$	Binary parameter of compliance with the maximum coverage distance for the assignment of a client $i$ to a WH at candidate location $j$ .
Decision variables	
$x_j$	Binary variable that establishes the opening of a WH in candidate location $j$ .
$y_{ij}$	Binary variable that establishes the allocation of attention of neighborhood $i$ to a WH located in candidate location $j$ .

The mathematical model formulated in Binary Integer Linear Programming is presented below with expressions (1) to (4).

$$\text{Min } Z = \sum_{j \in U} x_j \cdot CI_j + \sum_{i \in N} \sum_{j \in U} y_{ij} \cdot C_{ij} \cdot P_i \cdot d_{ij} \quad (1)$$

$$\sum_{j \in U} a_{ij} \cdot y_{ij} = 1 \quad \forall i \in N \quad (2)$$

$$\sum_{i \in N} a_{ij} \cdot x_j \geq \sum_{i \in N} y_{ij} \quad \forall j \in U \quad (3)$$

## IV. Conclusion

Through the application of this work, a proposal is defined that aims to improve the urban logistics of goods within the city of Medellín, specifically taking as an object of study the district 14, El Poblado. This proposal, as it has been seen in different reviewed works, is likely to generate significant contributions in several areas of the city, such as vehicular congestion and the generation of vehicular pollutants.

The development of the application of mathematical models to the realities of urban logistics can be considered a great tool when looking for improvements in this framework. These are easily adapted to the representation of real scenarios in cases throughout a city. From our point of view, it is a tool that generates great benefits and to which, nowadays, little visible use is given when making decisions in the framework of public sector administration.

Finally, as future work, three main complementary ideas can be identified. (i) Assess the capacities needed for the PM, to establish the space requirements to serve the assigned population. This can be done through design and sizing methods of logistics facilities. All this to optimize and achieve cost savings per facility. (ii) Depending on the demands and transports performed, an optimization of the resources allocated to the functions of picking and transport of orders could be applied. This to evaluate different work policies. For example, the transport personnel perform the picking in the warehouse, or on the contrary, separate the functions. I touch with a view to evaluate the need for hours allocated to each task and the need for handling equipment and transport of goods. Exact or heuristic optimization techniques can be used for this purpose. (iii) Complement the mathematical modeling to ensure that there is load balancing, so that the neighborhoods are distributed in a more equivalent way to the various WH proposed as a solution.

## Acknowledgment

This work was supported by Instituto Tecnológico Metropolitano (ITM) (Project no. P20239), in Medellín, Colombia.

## References

- [1] Burgos, G. (2021, January 13). Supply chain: Los retos del transporte ligero y la distribución urbana en tiempos del Coronavirus | América Retail. <https://www.america-retail.com/supply-chain/supply-chain-los-retos-del-transporte-ligero-y-la-distribucion-urbana-en-tiempos-del-coronavirus/>
- [2] Muñuzuri, J., Grosso, R., Escudero, A., & Cortés, P. (2017). Distribución de mercancías y desarrollo urbano sostenible. *Revista Transporte y Territorio*, 0(17), 34–58. <https://doi.org/10.34096/rtt.i17.3866>
- [3] Segura, V., Fuster, A., Antolín, F., Casellas, C., Payno, M., Grandío, A., Cagigós, A., & Muelas, M. (2020). Logística de última milla retos y soluciones en España. Deloitte. <https://www2.deloitte.com/content/dam/Deloitte/es/Document/s/operaciones/Deloitte-es-operaciones-last-mile.pdf>
- [4] Neira Marciales, L. (2020, March 27). “Durante la cuarentena por el virus Covid-19 se cuadruplican en el país los domicilios”. <https://www.larepublica.co/empresas/domicilios-se-cuadruplican-en-tiempos-de-cuarentena-por-el-covid-19-2983817>
- [5] Google Trends. Accessed February 2021. ([www.google.com/trends](http://www.google.com/trends)).

$$x_i, y_{ij} \in \{0,1\} \quad \forall i \in N, j \in U \quad (4)$$

The objective function (1) seeks to minimize the costs of the exercise. The first term represents the assumed cost for each facility. The second term that sums the transportation cost from the WH at location  $j$  to the assigned customer  $i$ . The group of equations (2) guarantees the assignment of each neighborhood  $i$  to an WH facility at a candidate location  $j$  complying with the maximum coverage limit. Equation group (3) guarantees that customers  $i$  are assigned to active WH at candidate location  $j$ . Finally, group of equations (4) ensures that the variables  $x_j, y_{ij}$  are of binary order.

## III. Results

The mathematical model has been implemented in Python 3.9 and solved with Cplex IDLE Studio 20.1.0 optimizer. The results have been found on a 2.3 GHz Ryzen 7 computer with 8 GB of RAM, running on Windows 10 Professional Operating System. The proposal seeks to establish a proposal for a warehouse management scheme from the public sector, with the purpose of private use, but with the aim of reducing trips for grocery services from distant distances in the city. Until to this point, we have started with these results by evaluating the model defined for the construction of the proposal at city level, we consider the results are efficient at this scale, given that we found a solution of optimal order and in an efficient way, with a run of 1 second of machine.

The results obtained propose the opening of 3 WH. The neighborhoods where the model yields the possible locations of the centers are: La Linde (WH3), Los Naranjos (WH5) and El Diamante No. 2 (WH7). The final assignment of each neighborhood to the various locations proposed for WH yielded a minimum cost objective function with a value of \$ 105,326,975.076 (\$COP) (\$ 28,833.30 US dollar). The results are plotted in Figure VI, the centroid points of the neighborhoods assigned to each of the WHs are highlighted by connecting lines reflected in different colors for each of them.

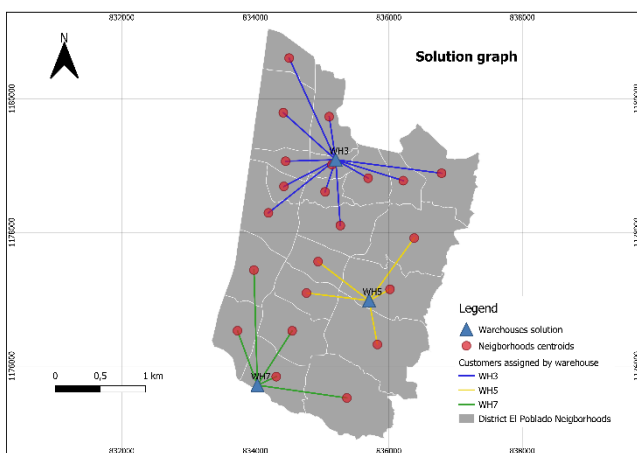


Figure VI. Solution graph.

- [6] Sopha, B. M., Sri Asih, A. M., Pradana, F. D., Gunawan, H. E., & Karuniawati, Y. (2016). "Urban distribution center location: Combination of spatial analysis and multi-objective mixed-integer linear programming". *International Journal of Engineering Business Management*, 8, 1–10. <https://doi.org/10.1177/1847979016678371>
- [7] Campos Magin, J. (2015, May). "Las plataformas logísticas de distribución urbana de mercancías: un elemento de desarrollo y regulación del transporte de mercancías en las ciudades". <https://upcommons.upc.edu/bitstream/handle/2117/27229/15572417.pdf>
- [8] de Carvalho, N. L., Vieira, J. G. V., da Fonseca, P. N., & Dulebenets, M. A. (2020). A multi-criteria structure for sustainable implementation of urban distribution centers in historical cities. *Sustainability* (Switzerland), 12(14). <https://doi.org/10.3390/su12145538>
- [9] Ariza Nieto, J. A. (2013). "Modelo de programación lineal basado en la caracterización de la cadena de suministro de los productos bovinos con alta producción en la provincia de sabana centro". *Journal of Chemical Information and Modeling*, 53(9), 1689–1699.
- [10] INRIX. (2020). 2020 Global Traffic Scorecard. INRIX Research. <https://inrix.com/scorecard/>
- [11] Alcaldía de Medellín. (2010). Primera Parte: Generalidades Medellín y su Población MEDELLÍN Y SU POBLACIÓN. <https://www.medellin.gov.co/irj/go/km/docs/wpcontent/Sites/Subportal%20del%20Ciudadano/Plan%20de%20Desarrollo/Secciones/Informaci%C3%B3n%20General/Documentos/PO T/medellinPoblacion.pdf>
- [12] DANE. (2019). Resultados Censo Nacional de Población y Vivienda 2018 (National population census). <https://www.dane.gov.co/index.php/servicios-al-ciudadano/60-espanol/demograficas/censos>
- [13] Alcaldía de Medellín. (2013). Documento de rendición de cuentas a la ciudadanía para la Comuna 14 El Poblado. Periódico Cuentas Claras, 1, 8. [https://www.medellin.gov.co/irj/go/km/docs/wpcontent/Sites/Subportal%20del%20Ciudadano/NuestroGobierno/Secciones/Plantillas/Genéricas/Documentos/2013/Cuentas Claras Comuna/1 octubre/comuna 14 baja.pdf](https://www.medellin.gov.co/irj/go/km/docs/wpcontent/Sites/Subportal%20del%20Ciudadano/NuestroGobierno/Secciones/Plantillas/Genéricas/Documentos/2013/Cuentas%20Claras%20Comuna/1octubre/comuna%2014%20baja.pdf)
- [14] Alcaldía de Medellín. (2016). Perfil Sociodemográfico por barrio Comuna 14 El Poblado 2016-2020. 223. <https://www.medellin.gov.co/irj/go/km/docs/pccdesign/Subportal%20del%20Ciudadano/2/PlandeDesarrollo/017/IndicadoresyEstadísticas/Shared%20Content/Documentos/ProyeccionPoblacion2016-2020/Perfil%20Demogr%C3%A1fico%20Barrios%202016%20%E2%80%93%202020%20Comuna%2014%20El%20Poblado.pdf>
- [15] QGIS Development Team Version 3.16.7-Hannover (2009). QGIS Geographic Information System. Open Source Geospatial Foundation, url: <http://qgis.osgeo.org>.
- [16] Rappi®. Accessed February 2021. [www.rappi.com.co/](http://www.rappi.com.co/).
- [17] Finca raíz. Accessed February 2021. <https://www.fincaraiz.com.co/>

# Close Price Prediction of Day Stock Markets with Machine Learning and NLP models

Purushoth Ananatharasa  
Informatics Institute of Technology  
anantharasa.2017152@iit.ac.lk

Ragu Sivaraman  
Informatics Institute of Technology  
ragu.s@iit.ac.lk

**Abstract**— The Stock market is an important factor that displays the development of a country. The reason for considering this statement is due to the transactions that happen in the market time brings out huge capital gain from investors and traders from trading with company shares. Therefore the transactions of investors and stock traders are very important to keep the market alive. Considering the stock traders in the short-term market, they analyze the performance of a company by the past values and performance of those company indexes. However, analyzing those close prices would not help. There are many systems that forecast the future prices for a long time in future but due to the high volatile nature, these values predicted from these systems cannot always be accurate. Therefore the limitation of analyzing both the past transactions and impacting features to predict the close price of a day shows the main research problem that should be addressed in the research. This trained solution used machine learning models such as Random Forest Regression, XGBoost, SVM and Lasso models to predict the next two days of the closing prices. Also an additional feature which analyses the news sentiment which affects the pattern of stock performance and prompts suggestions for traders. The evaluated system accuracy was measured with RMSE, MDA, MSE and the overall accuracy was up to 97% where the whole system was efficient and satisfactory during benchmarking with existing systems. Furthermore, the system was evaluated by domain experts and end users under well-designed evaluation criteria.

**Keywords**— *Stock Market Prediction, Machine Learning, Random Forest Regression, Sentiment Analysis.*

## I. Introduction

The stock market can be considered as an indicator of a country's economy since it acts as a platform of buying and selling stocks which includes many company indexes in sectors like agriculture, health, manufacturing etc, are being monitored by the public to invest capital [1]. Therefore a capital market acts as a transparent medium to evaluate the performance of the companies for the stock traders and the investors to trade-in. Stocks are considered as the most preferred trading medium where they can be categorized into long term and short-term stocks. The long-term stocks are considered as investments with a long time span of revenue generator and short-term stocks which generate income with a short time span to a daily stock trader [2]. This trade between the public sector and the companies provides many mutual benefits. The companies get benefits like investment gains where the capital which is required to expand the business is obtained from verified income channel. Diversification of value of the company into another category possibly brings the opportunities of revenues in the company profile [3]. From a stock trader's point of view, a person who has a company shares will have an income channel from the profit gained from a daily stock trade. Also, a shareholder will have the control of making decisions to buy or sell stocks specifically to the individual tolerance upon stocks.

A stock trader before purchasing stocks considers factors such as price patterns, trading pattern, opinions of public perception and the services offered by that specific company whereas, a stock investor consider other factors like financial summary, dividends, economic growth, cash flow etc. Therefore these both stock investors and traders are important factors for the stock market to function [4]. It is both the groups that keep the stock market transactions alive every day though throughout the past history. However there is never a solid price for a stock price for the traders to trade because the value starts fluctuating in the market making the values of the assets highly unpredictable. These fluctuations are based on capital flow in and out of financial reserve, also due to the competitiveness of the other companies in the domains [5]. This stature of the stock makes it highly volatile and unreliable. Considering huge companies like Apple (AAPL), it has many factors such as Apple events, product launches and many other internal company events to impact its stock performance in a day trade [6]. Therefore as a well experienced day trader it is important to analyze all the impacting factors like company events and past data up to a necessary time period before begin trading. In fact, there are prediction models which are used by the stock market experts to find the directions (up or down) for a specific number of days in future, but the systems cannot predict the closest stock price due to governing factors which makes the stock prices really challenging to predict and do daily trad.

Therefore a system which analyzing and predicting the closing price of certain companies for long time span has a possible consequence of misleading the stock traders to make incorrect decisions to do trade in non-ideal days, because predicting for a longer time span which is more than 5 days will have high chances of getting incorrect predictions due to the factors like market sentiments of customers and unexpected global events which could affect the whole industry. Therefore the existence of a system that analyzes the past data for a time span to predict the close price of the next 2 days (ideal) as the main function and provide the nature of the stock index (Ideal to trade or not) based on unexpected news published on the internet as another feature using machine learning is still a wish list system among the stock traders.

Therefore these below mentioned factors have to be considered when predicting the close prices of short term (daily trade) stock prices.

### A. Company News

The public opinion of a stock company is mainly portrayed by the news that is published in the public news media (positive/negative) [9]. This news published impacts on the investors, long term and short-term traders to evaluate the decisions which they made are wiser or profitable or non-profitable. It is the news which is delivered to the public that attracts the people to buy and sell or even invest in that specific company to perform any kind of capital transactions [10].

### B. Volume of stocks

The volume of trade is also another important factor for the stock traders to analyze the value of the stock index which has performed on a day which also indirectly used to determine the close price of a company [11]. The volume of trade depends on specific company. Generally if the volume of the stock is high, it directly brings out opinions that the close price might get a positive increment since a lot of customers have present in the trade to contribute to the overall volume to increase [9].

## II. Existing Works

Analyzing the existing works gives a deep understanding of the approaches and the system limitations that are in the present systems. After carrying out a comparison with each available system the close price prediction of a stock index can be done in three main approaches.

- Predicting close prices using past datasets.
- Predicting close prices using social media and financial news.
- Predicting close prices using Technical Indicators.

### A. Predicting close prices using past datasets

Using ensemble approach like Random Forest can be used to predict the stock price for specific days and compared with a deep learning model for the obtained values in this research [13]. In this approach, the model used input parameters such as Open, High, Low and Close prices for prediction. The model was able to provide results from the Random Forest model where the model was validated with evaluation metrics such as root mean squared error (RMSE), Mean Absolute Percentage Error (MAPE) and Mean Bias Error (MBE) for the unclassified data for five distinct instances.

However this novel approach was challenged from [14] where LASSO and Ridge models were used to predict the close price. The obtained results were same enough with the existing works to benchmark. To evaluate these models same metrics were used (MAPE and RMSE). The following table explains the critiquing of the previous approaches and systems that were completed with the past data prediction approach with the features and limitations that were present.

Table I. Existing work on past datasets.

Research	Features	Limitation
[13]	Predicts the price for one day.	Higher RMSE, MAPE values compared to other models.
[15]	Displays result up or down by analyzing the past datasets	The reinvention of traditional model with more beneficial functions
[16]	Obtained 70% accuracy for the model.	Need to add sentiment data due to impermanent nature of close price.
[14]	Used LASSO and Ridge model to predict company stocks	Did not consider any market impacting factors or market policies.
[17]	Gave acceptable results compared to ANN and ARIMA models.	Predicted scores were affected by trader sentiment.
[18]	Used multiple ML models (XG Boost, SVC, KNN, ANN)	Deep Learning models had better confusion matrix
[19]	The prediction was done with granger causal factor between input parameters.	Limited performance.

### B. Predicting close prices using social media and financial news

Social media impacts the decision of day traders involving in the trade. Specially to analyses whether it is profitable to perform trade in the market. This impacts the overall performance of the market at the close time. To examine this factor the author [12] carried a research of analyzing the social media feeds and other public news with models like SVM, KNN, NN and Linear Regression. In summary with other researches the positive and negative news affected considerably the performance of the closing price. Therefore it was decided to analyze both factors to find the direction and value of the close price. This table explains the critiquing of the previous researches and systems that were completed with the social media approach with the features and limitations that are present.

Table II. Existing work on social media and news articles.

Research	Features	Limitations
[20]	Used SVM and RF classifiers	Add more companies for prediction model and use the same technique for algorithmic trading.
[12]	Used Random Forest, SVM, kernel factory and AdaBoost algorithms	Models are made based on certain regional markets
[21]	Accuracy is 56.07%. Can do prediction for large scale datasets.	Number of topics and sentiment must specify.
[22]	Textual representation is better than numerical data sets and bag of words methods were sued.	For the best profit-making qualities technical indexes such as MA and MACD has to be included.

### C. Predicting close prices using Technical Indicators

In the research of [15], the team decided to use the tree-based classifiers approach to predict the stock prices of certain companies go up or down (classification mechanism) for a specific number of days in future, based on stock market indicators such as **MACD** and **RSI**. At the end of prediction, the Random Forest classifier had better accuracy than the Gradient Boost Algorithm to the selected companies. However, the team also noticed that the F1 score of both models increased when they increased the window width of the model. Since this model analyzes all these indicators the author had to find correlation between these input parameters to predict the output [15].

In another research of Condo Kim, the author used Effective transfer entropy (**ETE**) which related to Granger casual relationships (analyze daily price, exchange price and interest rates) to predict the stock price [19]. For this research ML models such as XGBoost, Linear regression and RF were used with LSTM model. Among all the other models RF and LSTM models predicted the future pattern of stocks (long term) with less noise in the data flow compared to XGBoost and linear Regression. The author mentioned that the models have predicted the direction of stock price based on the ETE might be affected by the financial events based on US Stocks which made a conclusion to the price direction might be impacted due to any sudden reasons in the US Stock market.

Considering the past works, researches have been done on both classification and regression predictions to predict the stock prices depending on this scenario. In research of [9], the author has predicted the average price of certain companies for the next (3-10) days by using SVR and kernel of Radial Basis Function (RBF). In this research, the author has compared the obtained accuracy of the certain companies obtained through the model where the technical indicators were given as input parameters to predict the close price. However, the author has highlighted the obtained results might be changed since the timespan of closing price is more than 3 days and market sentiments could impact in the predicted prices.

### III. Methodology

From the above discussions it was decided to go with the historical transactions and social media, other financial news which leverages on prediction of close price. Therefore utilizing suitable models to find the close price using ensemble method and nature of the stock index based on the news is focused in this section. The product consists two main components which are “Price Prediction model”, “News Analyzer model” to predict the price and nature of a stock index. For this scenario the Apple stock index (**AAPL**) and Dow Jones (**DWJA**) were chosen.

### A. Price Prediction Model

In this component an efficient model is utilized among other models by inserting a valid dataset that are fed into the regression models. The training datasets were obtained from Yahoo finance. Before training the model the imported datasets were pre-processed with missing values, validated with datasets without noise in the imported data.

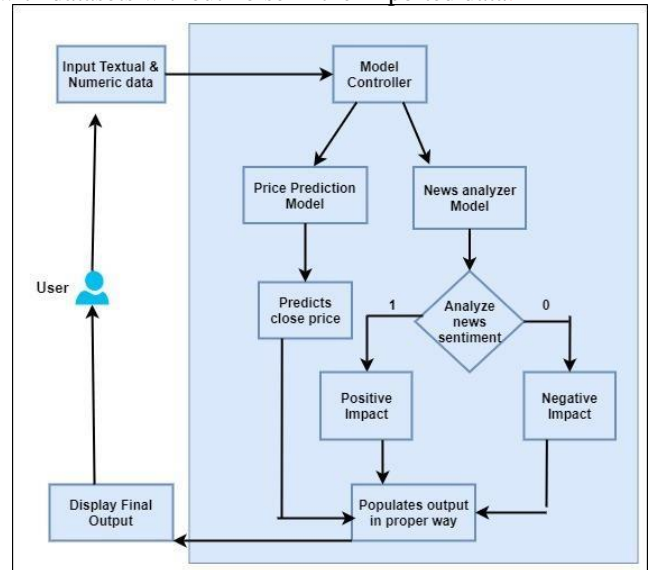


Figure I : Overview Diagram of the prediction system

Before training the model, it was decided to input the open, high, low and volume columns of the dataset to predict the close price from this regression model. The Figure II depicts the requirement gathered from the end users and domain experts regarding the core functionalities during the system requirement gathering phase. Finally after encoding and feature scaling the datasets were split with 80%-20% for training and testing. In the training phase suitable parameters were identified by grid search importing from scikit library. The parameters such as criterion, n\_estimators, verbose and max\_iter were recommended most suitable for these models.

3). Are you aware of Open, close, high, low price of stocks and other basic concepts?  
109 responses

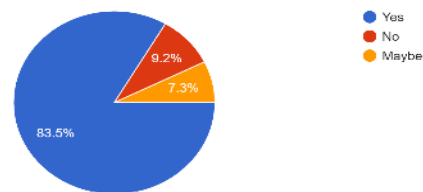


Figure II : Survey from end users on most used data

### B. News Analyzer Model

This model was based on NLP which associates with human languages in the form of audio or textual information. In this scenario which involves textual information, sentiment analysis was used for classification. This model has news datasets that were obtained from trusted news sources like verified users in twitter, public news figures that provide financial news on capital market. This dataset was obtained from Kaggle, an open source data platform for data science and machine learning projects. The imported data was done a pre-processing step which consists of removing English stop words and unwanted characters that could affect

the training process. Following that tokenizing and stemming techniques were carried out to reduce the noise among the datasets. After benchmarking with multiple models for better accuracy the SoftMax logistic model was selected to classify the news whether it could affect the stock price positively or negatively.

The overall accuracy of the model was calculated by the confusion matrix which consists of components like True positive, true negative, false positive and false negative. Other additional metrics such as Recall , F1, Precision Score were also calculated using these formulas.

Table IV. Time consumed for Model Training

Model	Time Taken (Seconds)
Random Forest Regressor	31
XG Boost	48
Decision Tree	66
LASSO	55
SVM	51

In the same criteria the Random Forest regression model trained with the least amount of time with the most accurate value compared to the others with the necessary input parameters. Finally to validate the model, K-Fold validation was carried out to validate whether the model is overfitting or underfitting. And the table below shows the comparison of actual predicted value for randomly taken for a sample of three days, which was convincing for suggestion and proved the model is balanced and can be reliable for critical prediction purposes

Table V. Actual vs Predicted closing prices (AAPL).

Date	Real Value (Yahoo Finance.com)	Predicted Values
29 <sup>th</sup> April 2021	133.48	133.62
13 <sup>th</sup> May 2021	124.97	124.11
26 <sup>th</sup> May 2021	126.85	126.89

To compare the results obtained, a benchmarking process was carried out with the existing work in order to validate the price prediction model has comparatively produced better results. The table below explains the benchmarking process.

Table VI. Benchmarking of system with existing system.

Loss functions	Our system	[13]
RMSE	0.37	0.42
MAPE	0.96	0.77
MAE	0.13	N/A

According to this table this shows the accuracy of the classification model (News Analyzer Model) for impacting on the performance of the stock index. From the observation the SoftMax logistic model gave the highest classification accuracy around 97% after the training phase. The results of the model was obtained after training the dataset with cleaned data which consisted of neutral and non-ideal news and comments in the news sections.

Table VII. Accuracy of Classification model

Classification Model	Accuracy (%)
Random Forest	96.63
SoftMax Logistic	97.07
Multinomial Naïve Bayes	93.00
Bernoulli Naïve Bayes	95.03

Furthermore in order to support the results obtained, the heatmap and the roc curve of the SoftMax model is presented below

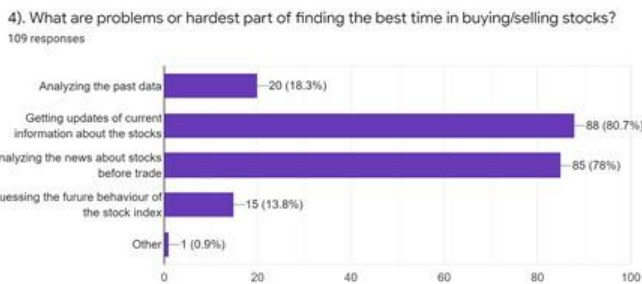


Figure III : Survey from end users on most used medium

#### IV. Discussion and Results

According to this section, the experimental results obtained were discussed here. The main purpose of this research is to predict the close price of the stock index with improved accuracy with the impact of the news or social media on the performance of the stock index. Furthermore the implemented prototype was evaluated with evaluation criteria in order to make sure it has acquired the required functionalities that are expected from the end user expectations.

As mentioned earlier, the aim is to design separate models to predict the price for today and predict the impact of news on the performance of the stock. Therefore the table below shows the results of loss function (MAPE, RMSE, MSE) obtained from trying out different machine learning models with suitable input parameters which are decided from gridSerachCV method.

Table III. Loss function values of Price Prediction Models

Regression Model	MAE (%)	MAPE (%)	RMSE (%)
Random Forest	0.13	0.96	0.37
XG Boost	0.13	1.20	0.34
Decision Tree	0.18	1.24	0.50
SVM	0.26	2.34	0.75
LASSO	0.19	5.87	0.40

From the observation the random forest regression model out performed among the other regression models with less loss function score for the same instances. Also the compiling time of the models were also measured in order to find an effective model to process the dataset with less amount of time. The table below describes the compilation time taken to train the regression models for predicting the close price.

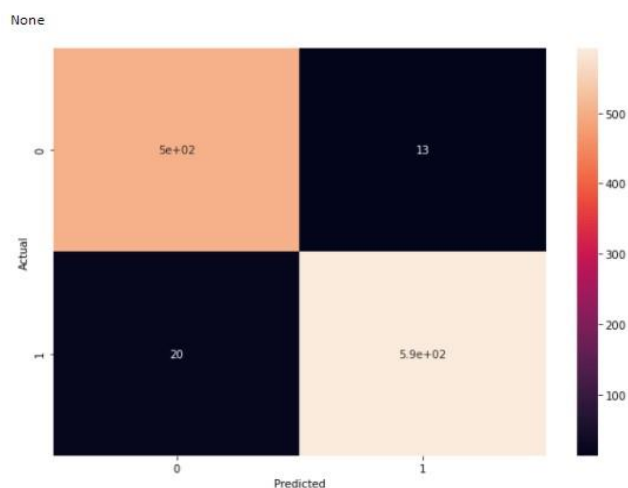


Figure IV. Heatmap of the SoftMax model

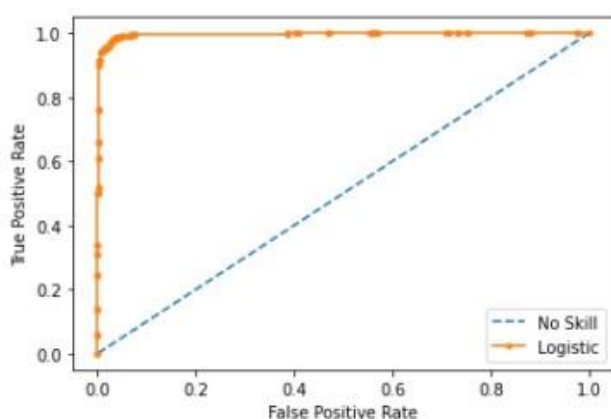


Figure V: ROC curve of the SoftMax model

## V. Conclusion

A gap for an accurate and effective system for a suggestion system based on the past data and public perception still exists in the market. This research paper discusses a novel approach to get the most accurate close price of the next two days based on the historical datasets and the impact of the public news on the stock index performance. The random forest regression model to predict the price of the stock using machine learning was selected after comparing benchmarking the models. The SoftMax logistic model was designed to analyze the impact of the public news headlines or feeds posted on the social media. Both the models have been trained with most ideal input parameters and trained under most common scenarios to perform as ideal models.

For the time being, the system provides predictions for only two stock indices AAPL and DWJA since those indices have high volatile character due to many global and economic uncertainties. For future enhancements more stock indexes will be added to the prediction system based on the different market regions based on the locations. Another functionality will be adding more stemming techniques to the News Analyzer Model. Stemming techniques may vary based on the found datasets such as Lancaster Stemmer and Lemmatization etc for data cleaning.

Adding vectorization techniques to test exception scenarios in the trading market is also another enhancement that will be added. Finally adding the feature of analyzing the technical indicators such as Moving Average, Fibonacci average to predict the movement of stock for far long time period is also to be another functionality which could ultimately addresses the issues faced by the stock market experts and end users with the misconceptions and unexpected errors in the capital markets.

## References

- [1] Abraham, A., Krömer, P. and Snášel, V. (2015) 'Afro-European Conference for Industrial Advancement: Proceedings of the First International Afro-European Conference for Industrial Advancement AECIA 2014', *Advances in Intelligent Systems and Computing*, 334, pp. 371–381. doi: 10.1007/978-3-319-13572-4.
- [2] Akita, R. (2016) 'Deep learning for Stock Prediction Using Numerical and Textual Information', *2016 IEEE/ACIS 15th International Conference on Computer and Information Science (ICIS)*, pp. 1–6. doi: 10.1109/ICIS.2016.7550882.
- [3] Al-Jaifi, H. A. (2017) 'Ownership concentration, earnings management and stock market liquidity: evidence from Malaysia', *Corporate Governance (Bingley)*, 17(3), pp. 490–510. doi: 10.1108/CG-06-2016-0139.
- [4] Ballings, M. *et al.* (2015) 'Expert Systems with Applications Evaluating multiple classifiers for stock price direction prediction', *EXPERT SYSTEMS WITH APPLICATIONS*. Elsevier Ltd, (May). doi: 10.1016/j.eswa.2015.05.013.
- [5] Basak, S. *et al.* (2019) 'Predicting the direction of stock market prices using tree-based classifiers', *North American Journal of Economics and Finance*. Elsevier, 47(December 2017), pp. 552–567. doi: 10.1016/j.najef.2018.06.013.
- [6] Cakra, Y. E. and Distiawan Trisedya, B. (2016) 'Stock price prediction using linear regression based on sentiment analysis', *ICACSIS 2015 - 2015 International Conference on Advanced Computer Science and Information Systems, Proceedings*, pp. 147–154. doi: 10.1109/ICACSIS.2015.7415179.
- [7] Cervelló-Royo, R., Guijarro, F. and Michniuk, K. (2015) 'Stock market trading rule based on pattern recognition and technical analysis: Forecasting the DJIA index with intraday data', *Expert Systems with Applications*, 42(14), pp. 5963–5975. doi: 10.1016/j.eswa.2015.03.017.
- [8] Clark, E. and Kassimatis, K. (2017) 'Country financial risk and stock market performance: The case of latin america', *Evaluating Country Risks for International Investments: Tools, Techniques and Applications*, 56(1), pp. 117–148. doi: 10.1142/9789813224940\_0005.
- [9] Di, X. (2014) 'Stock Trend Prediction with Technical Indicators using SVM', *Stanford University*.
- [10] Gaillard, P. (2004) 'Rwanda 1994: "...kill as many people as you want, you cannot kill their memory"', *International Committee of the Red Cross*, pp. 1–24. doi: 10.6084/m9.figshare.5028110.
- [11] Greenwald, D., Lettau, M. and Ludvigson, S. (2014) 'Origins of Stock Market Fluctuations', *NBER Working Paper Series*, 19818. Available at: <http://www.nber.org/papers/w19818.pdf>.
- [12] Jin, Z. *et al.* (2020) 'The industrial asymmetry of the stock price prediction with investor sentiment: Based on the comparison of predictive effects with SVR', *Journal of Forecasting*, 39(7), pp. 1166–1178. doi: 10.1002/for.2681.
- [13] Joshi, K., H. N. B. and Rao, J. (2016) 'Stock Trend Prediction Using News Sentiment Analysis', *International Journal of Computer Science and Information Technology*, 8(3), pp. 67–76. doi: 10.5121/ijcsit.2016.8306.
- [14] Kim, S. *et al.* (2020) 'Predicting the Direction of US Stock Prices Using Effective Transfer Entropy and Machine Learning



- Techniques', *IEEE Access*, 8, pp. 111660–111682. doi: 10.1109/ACCESS.2020.3002174.
- [15] Maio, P. and Santa-Clara, P. (2017) *Short-Term Interest Rates and Stock Market Anomalies*, *Journal of Financial and Quantitative Analysis*. doi: 10.1017/S002210901700028X.
- [16] Nabipour, M. *et al.* (2020) 'Predicting Stock Market Trends Using Machine Learning and Deep Learning Algorithms Via Continuous and Binary Data; A Comparative Analysis', *IEEE Access*, 8, pp. 150199–150212. doi: 10.1109/ACCESS.2020.3015966.
- [17] Nguyen, T. H. and Shirai, K. (2015) 'Topic modeling based sentiment analysis on social media for stock market prediction', *ACL-IJCNLP 2015 - 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing of the Asian Federation of Natural Language Processing, Proceedings of the Conference*, 1, pp. 1354–1364. doi: 10.3115/v1/p15-1131.
- [18] Pradhan, R. S. and Dahal, S. (2018) 'Factors Affecting the Share Price: Evidence from Nepalese Commercial Banks', *SSRN Electronic Journal*, pp. 1–16. doi: 10.2139/ssrn.2793469.
- [19] Shah, D., Isah, H. and Zulkernine, F. (2018) 'Predicting the effects of news sentiments on the stock market', *arXiv*, (1), pp. 1–4.
- [20] Skuza, M. and Romanowski, A. (2015) 'Sentiment analysis of Twitter data within big data distributed environment for stock prediction', *Proceedings of the 2015 Federated Conference on Computer Science and Information Systems, FedCSIS 2015*, 5, pp. 1349–1354. doi: 10.15439/2015F230.
- [21] Victor Chow, K. *et al.* (1995) 'Long-term and short-term price memory in the stock market', *Economics Letters*, 49(3), pp. 287–293. doi: 10.1016/0165-1765(95)00690-H.
- [22] Vijh, M. *et al.* (2020) 'Stock Closing Price Prediction using Machine Learning Techniques', *Procedia Computer Science*. Elsevier B.V., 167(2019), pp. 599–606. doi: 10.1016/j.procs.2020.03.3

# Investigation of Permeability Coefficient in Layered Soils

Kaveh Dehghanian  
Department of Civil Engineering  
Istanbul Aydin University  
Istanbul, Turkey  
kavehdehghanian@iau.edu.tr

Mohammad Haroon Saeedi  
Department of Civil Engineering  
Istanbul Aydin University  
Istanbul, Turkey  
haroonsaidi64@gmail.com

**Abstract**— The hydraulic conductivity of permeable media is a critical property that depends upon different properties of soil mass such as porosity, size, and shape of soil particles, initial water content, and compaction. As the characteristic condition, the soil mass exists in layered strata, hence it is called as stratified soil. Soils are permeable materials due to the presence of interconnecting spaces that enable fluids to flow when there is a difference in energy head. The shape and size of particles, in turn, affect the interconnecting voids. Water flow through a soil mass, is proportional to the size of the void apertures rather than the overall number of voids. Even though void ratios are frequently greater than for fine grained soil. The relative position and thickness of a soil layer in a stratified soil system, are two critical variables that determine the permeability of the composite soil layer. In this study, a series of falling head tests were performed to determine the permeability of two-layered soils using two types of bentonite and sand, as well as the Atterberg limit, sieve analysis, specific gravity, and proctor test. It is shown that increase in Atterberg limits results in decrease of permeability. The higher the specific gravity, the higher the permeability coefficient. The permeability is higher at higher void ratio. Furthermore, the permeability of stratified soils is affected by the thickness of the end layer.

**Keywords**—Permeability Coefficient, Layered Soil Profile, Falling Head Test, Atterberg Limit, Void Ratio

## I. Introduction

In classical soil mechanics, soil is considered as a homogeneous and isotropic material. In most of the cases, the experiments and numerical analysis are performed for a single layer while the soil is a layered medium in the field. The permeability coefficient is often obtained by a constant head permeability test for coarse-grained soils and a falling head test in fine-grained soils. The assessment of permeability is significant for erosion control, slope stability control, wastewater management, and structural failure due to foundation settlement problems. For layered soil systems, the layers can either be horizontal, vertical, or inclined. Each level has its own permeability coefficient,  $k$ . The typical or equivalent permeability coefficient of the stratified deposit,  $k_{eq}$ , is total of the direction of flow in respect to the orientation of the bedding planes. The coefficient of permeability ( $k$ ) of soil masses is calculated using Darcy's law. When the flow is parallel to the bedding planes' orientation, the equivalent coefficient of permeability of a stratified soil deposit is derived by

$$k_{eq} = \frac{\sum_{i=1}^n L_i}{\sum_{i=1}^n \left(\frac{L_i}{k_i}\right)} \quad (1)$$

Where  $L_i$  is the thickness of the  $i$ th layer in the layered profile and  $k_i$  is the coefficient of permeability of that layer. The permeability coefficient of the bentonite clays is quite low and is traditionally measured by falling head permeability test. This method gives results over a long period of time as the sample is expected to saturate [1].

"Uppot et al. (1989) investigated two clays subjected to organic and inorganic permeants to study the changes in permeability caused by the reaction between clays and permeants [2]. Afterward, Haug et al. (1990) a prototype liner formed of Ottawa sand and sodium bentonite. This material was mixed, moisture-conditioned, and compacted into reinforced wooden frames. The in situ permeability test results were verified with low gradient, back-pressure saturated triaxial permeameter tests conducted on undisturbed cored and remolded samples [3]" . "Sridharan and Prakash (2002) researched on two-layer soil frameworks demonstrates that the shared interaction among distinctive layers of diverse soil sorts shaping a stratified store influences the proportionate permeability of the stratified store, which cannot be essentially calculated by the utilize of the condition for the proportionate coefficient of the porousness of a stratified store when the stream is typical to the introduction of the bedding planes based on the Darcy's law. The porousness of the exit layer controls whether the measured porousness is more prominent or lesser than the hypothetical values for a stratified store" [4].

"Galvaeo et al. (2004) performed another test in which coefficient of permeability of saprolitic soil increased about five times when two percent lime was added and then decreased on further addition of lime. This is assign to the creation of chemical bonds and aggregation. As for lateritic soil, the coefficient of permeability decreased as lime was added. This is also assign to the same mechanism except that the bonds are weaker than those developed in Soil [5]. Nikraz et al. (2011) carried out a series of laboratory permeability tests to evaluate fiber effect on hydraulic conductivity behavior of composite sand. Clayey sand was selected as soil part of the composite and natural fiber was used as reinforcement [6]. Sridharan and Prakash (2013) conducted a comparative study of the measured equivalent coefficient of permeability of three-layer soil sediments with the theoretically calculated values has been made. The results demonstrate that, by and large, the coefficient of permeability of the bottom layer controls whether the measured value of equivalent coefficient of permeability is greater or lesser than the theoretically calculated value also when a stratified soil deposit contains more than 3 layers, different combinations of positioning of layers of different  $k$  values are possible. Hence, in such cases, it becomes difficult to predict whether the measured value of  $k_{eq}$  is less or equal to or more than that calculated. [7]" . The consequence of this observation is the realization that the equivalent coefficient of permeability of any layered soil deposit is not just dependent upon the values of  $k$  of the individual layers constituting the deposit, and that it also depends upon the relative positioning of the layers in the system. Sridharan and Prakash (2002) studied two-layer soil systems with equal thickness layers, whereas in the current research, sand and bentonite were treated in varied thickness layers with varied sample sizes. The main goal of this paper is to measure and compare various sizes of sand and bentonite in soil layers.

## II. Experimental study

Within the scope of the article, soil samples with varied proportions were employed. Each soil sample's studies are reported separately and in chronological sequence. Bentonite clay has a delicate, silky texture and is a natural clay. When combined with water, it makes a paste. Elito Bentonite Clay manufactured from a district in Izmir, southwestern of Turkey and sand from concrete plants and an area close to the university campus. The tested samples were selected in the following percentages: 20% clay bentonite +80% sand, 30% clay bentonite+70% sand, and 40% clay bentonite + 60% sand. "The uniformity coefficient (Cu) indicates the variance in particle sizes in soil and is defined as the ratio of D60 to D10. D60 denotes the grain diameter at which 60% of soil particles are finer and 40% are coarser, whereas D10 denotes the grain diameter at which 10% of particles are finer and 90% are coarser" [8]. Table I depicts the values of  $C_u$  and  $C_c$  for the samples.

"Atterberg limits were performed according to these methods are still being used to determine the Liquid Limit, Plastic Limit and Shrinkage Limit of soils, which are outlined in ASTM D4318 and TS 1900-1" [8]. Table II depicts the values of Plastic Limit (PL), Liquid Limit (LL), Plasticity Index (PI) and shrinkage Limit for the soil samples.

Table I: Determination of CC, CU

RATIOS	$C_u$	$C_c$
100% sand	3.22	0.825
80% sand + 20% bentonite	3.928	0.982
70% sand + 30% bentonite	4.545	1.223
60% sand + 40% bentonite	4.85	0.945

Table II: Atterberg Limits test result

RATIOS	100% bentonite	80% sand+20% bentonite	70% sand+30% bentonite	60% sand+40% bentonite
Liquid Limit	167.71	26.44	43.39	55.01
Plastic Limit	50.53	NP	NP	NP
Plasticity Index	117.18	26.44	43.39	55.01
Shrinkage Limit	14.17	3.63	3.665	5.01

In Atterberg tests, bentonite has the highest value, as seen in the table II. The liquid limit reduces by the decrease in the bentonite percentage, depending on the qualities of bentonite soils, whereas the liquid limit increases as the bentonite ratios increase. Bentonite sample, which has less voids, necessitates a large amount of water to be saturated since bentonite has less voids due to its grains and features, therefore the result demonstrates that the liquid limit is extremely high. On the other hand, when the grain of the sand sample mixed with the bentonite its look like void of sand filling with bentonite throw the time therefore, as the gaps increase, the liquid limit decreases. The sample is not reduced to 3mm in diameter throughout the test in the case of composite bentonite and sand since the ratio of sand is likely larger than bentonite, therefore the plastic limit is zero. The sample with 25 impacts will be

left in the oven for 24 hours to examine how much it shrinks as the temperature rises, the more sand the lower shrinkage limit becomes.

Specific gravity tests are also done based on TS 1900-1 standards. Specific gravity of the combinations are as follow:

$$60\% \text{ sand} + 40\% \text{ Bentonit} = 2.521$$

$$70\% \text{ sand} + 30\% \text{ Bentonit} = 2.548$$

$$80\% \text{ sand} + 20\% \text{ Bentonit} = 2.553$$

$$100\% \text{ Bentonit} = 2.318$$

$$100\% \text{ sand} = 2.588$$

According to the specific gravity test results, the 100% sand has the highest specificity of the test, and the 100% bentonite clay sample has the lowest specific gravity of the test. As the amount of sand in the composite sample increases, specific gravity of sample increases too. Before beginning the falling head permeability test, the amount of water added to each test must be considered carefully. Proctor test defines how much water should be added to bentonite and sandy soils directly. The results of proctor test are shown in Table III.

Falling head test results are performed based on CEN ISO/ TS 17892-11 standard and are shown in table IV.

Table III: Determination amount of water and specific gravity test

Sample weight (gr)	Sample ratios	Water amount	Specific gravity
2110	100% B	0.45	2.31
1805	100% sand	0.17	2.58
1805	80% S+20% B	0.2	2.55
1961	70% S+30% B	0.3	2.561
2079	60% S+40% B	0.35	2.57

Table IV: permeability result for mixed samples.

RATIO	100%S	10% B	90%S+10%B	80%S+20%B	70%S+30%B	60%S+40%B
K (cm/s)	$2.8 \times 10^{-4}$	-	$3.2 \times 10^{-7}$	$4.25 \times 10^{-6}$	$1.4 \times 10^{-10}$	$5.7 \times 10^{-8}$

The permeability of 100 percent sand has the highest permissibility in the test, while the mixed sample has the lowest permissibility of 40 percent bentonite.

As shown in the table IV, as the amount of bentonite in the experiment drops, the amount of sand increases directly because the sand cavity is filled with bentonite. We might also infer that in the test, bentonite which is the fine grained material has low permeability due to the high plasticity and swelling capability of these soil types. It was observed that 100 percent bentonite did not saturate after 3 months, because this soil has a strong adhesion and water goes through the pore slowly, it absorbs water and keeps it from passing while in the beginning the water going through the sample is so fast. We attempt to determine the porosity and void ratio of the same sample after calculating the permeability of composite (sand, bentonite).

Table V: void ratio (e) and porosity (n) of different samples.

Weight (gr)	Ratio	Wet unit weight (gr/cm <sup>3</sup> )	Dry unit weight (gr/cm <sup>3</sup> )	(e)	(n)
2110	100%B	1.892	1.304	0.771	0.435
1805	100%S	1.618	1.383	0.86	0.462
1805	80S+20 B	1.618	1.348	0.892	0.471
1961	70S+30B	1.752	1.344	0.888	0.470
2079	60S+40B	1.864	1.380	0.862	0.462

The sample with 100% bentonite clay has the lowest void ratio and 20% bentonite has the highest void ratio relative to the test sample. As the ratio of bentonite increases, the void ratio should decrease, on the contrary, as the sand increases, the void ratio should also increase.

The method for testing the sample as 3 layers is the same as 2 layers, by putting the sample in 3 different layers as bentonite in the bottom of the mold. In the case that bentonite is placed at the top and sand at the bottom, the sand sample is discarded during the compression of the bentonite. The mold is separated into three layers, with filter sheets provided to prevent the production of mixed layers between the bottom and middle layers and the middle and top layers. The drop height permeability test was done after establishing a steady-state flow condition under the imposed maximum hydraulic gradient in all cases.

The tests were repeated with a new bentonite sample to see the impact of soil minerals and its physical and mechanical properties on the permeability. The optimum value of the test is 60% as the graph directly shows the maximum optimum of the proctor test taking place in the 60% bentonite zone as it will rise to 1970 kN/m<sup>3</sup> and decrease after 1970 kN/m<sup>3</sup>. Therefore, decide that the water added to this bentonite sample is 60 percent by weight of the bentonite sample. Therefore, the previous bentonite clay sample differs from this type because the amount of water added to the old sample was found to be 45 percent. However, when the grains of this sample approach a finer or even powder-like sample, it is expected that the amount of added water, which makes the sample saturated, is high.

After finding the bentonite optimum water content and Atterberg limits excluding the shrinkage limit for the new sample, permeabilities of the samples are achieved in two different methods: stratified and mixed states, as was done in earlier testing. The bentonite and sand ratios will be different from previous sample, so 5%, 10% and 15% are bentonite and 95%, 90% and 85% are sand sample amounts.

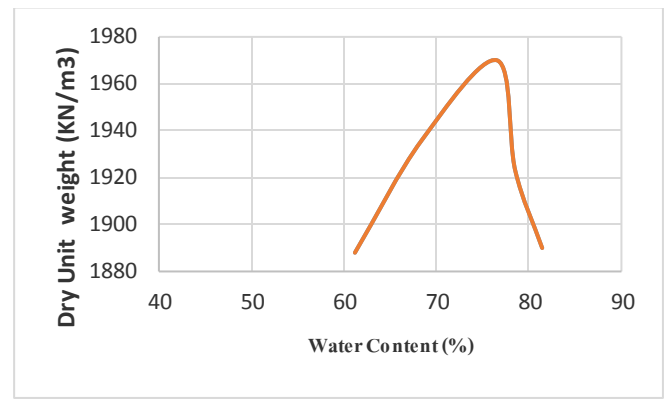


Figure I: dry unit weight versus water content

In this experimental study, sand and bentonite layers are used to create a stratified soil. After the requisite compression, samples are made by adjusting the thickness of each sample and filling them into three layers. Since the initial bentonite thickness is 0.7 cm and is too thin to be compacted, alternate rammers are used to compact the primary bentonite stratified soils.

1. Soil samples are obtained in the laboratory in brass pipes with a diameter of 4.8 mm, 1.084 mm and 10.84 mm and a length of 151.5 mm, or in larger containers from which the samples will be trimmed.
2. Take 2200 gr sand and 250 gr bentonite representative soil and mix with water if necessary to obtain.
3. After the soil sample is placed in the permeability mold, the saturation process should be continued until it is 100% saturated with water. The result for layered soils will be shown in the table below.

Table VI: Permeability test results for new stratified soil sample.

SAMPLE RATIO	layer Thickness (cm)	Permeability (cm/s)	
		Average Result	
(95%S+5%B)	B = 0.7 S = 13.4	1.75E-04	
(90%S+10%B)	B = 1.4 S = 12.7	1.68E-04	
(85%S+15%B)	B = 2.12 S = 12	1.10E-04	

The permeability coefficients of both layers composing the layered deposit must be different from their individual values when analyzed individually because the  $K_{eq}$  value is between  $K_{inlet}$  and  $K_{exit}$  and the continuity of flow must be assured across the thickness of the layered bed. Each layer in a soil profile might have similar or dissimilar qualities to the one above or below it.

"The individual permeability of these three sediments were computed directly from the equation for falling head permeability testing process since the validity of Darcy's law was proven for the soft sediments of the three soils at relatively low hydraulic gradients" [4].

$$k = 2.303 \frac{a_{in} a_{out} L}{At(a_{in} + a_{out})} \log_{10} \left( \frac{h_1}{h_2} \right) \quad (2)$$

"Where  $a_{in}$  = cross sectional area of the reservoir containing the influent liquid;  $a_{out}$  = cross sectional area of the reservoir containing the effluent liquid;  $A$  = cross sectional area of the specimen;  $t$  = elapsed time between the determination of  $h_1$  and  $h_2$ ;  $h_1$  = head loss across the specimen at time  $t_1$ ;  $h_2$  = head loss across the specimen at time  $t_2$ " [7].

Table VII summarizes the results of the measured and theoretical values. In most situations, the observed permeability values exceed the theoretical values, as seen in this table. This is due to the fact that bentonite and sand have different individual permeability. It's also worth noting that the composite layer's measured permeability is influenced by the composition of the exit layer. The composite layer with the bentonite layer at the exit has a lower equivalent permeability than the sandy composite layer at the exit.

Table VII: Measured and Theoretical permeability calculation.

Sample	Measured permeability	Theoretical permeability
(95%S) (top) (5%B)	8.75E-05	8.72 E-05
(90%S) (top) (10%B)	8.53E-05	8.50 E-05
(85%S) (top) (15%B)	5.59E-05	5.50 E-05

With bentonite at the top and sand at the bottom, the measured permeability decreases with increasing thickness of the top layer. The reluctance of bentonite soil to flowing through interconnected areas as the amount of water flowing through it decreases is the cause for this downward tendency. On the other hand, as the thickness of the top layer increases, the permeability tested with sand on top and bentonite on the bottom increases. Due to the increased load produced by sand on bentonite, sand has a substantially higher permeability than bentonite. The procedure for the composite sample is the same as before, but slightly different, with some details. The new amounts of the test procedure are 5% b+95% S, 10% B+90% S, and 15% B+85% S. The optimum of bentonite increased from 45 percent to 60 percent when greater than the previous sample. However, the optimum water of the sand is 17 percent of its weight as before.

Take 2200 g sand, 110 g, 220 g, 330 g bentonite representative soil and mix with water if necessary.

- For 100% sand: 17% weight of the sand sample should be adding water.
- For 100% bentonite clay: 60% weight of the bentonite clay sample should be adding water.
- For 5% bentonite clay + 95% sand: 66 g + 374 g water should be added.
- For 10% bentonite clay + 90% sand: 132 g + 374 g water should be added.
- For 15% bentonite clay + 85% sand: 198 g + 374 g water should be added.

### III. CONCLUSION AND DISCUSSION

In the laboratory, different types of permeability tests were carried out using bentonite and sandy varying proportions soil, and the equivalent permeability coefficient was determined for two types of layered soils.

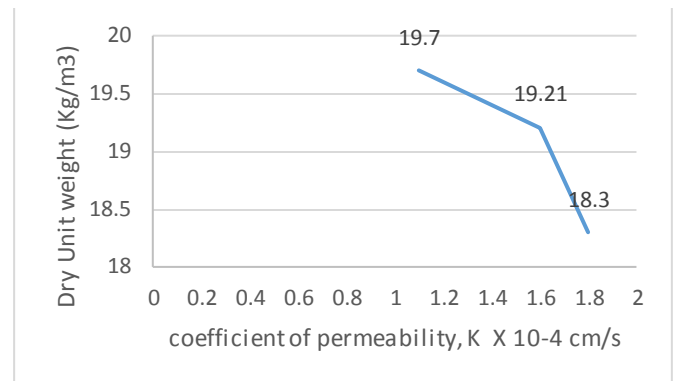


Figure II: Coefficient of permeability and dry density (newsample).

Table VIII:  $k$  and  $\gamma_d$  for proposed sample.

Ratio	$k$ (cm/s)	$\gamma_d$ (gr/cm <sup>3</sup> )
80% S+20%B	1.73E-06	1.4
70% S+30%B	1.91E-07	1.38
60% S+40%B	5.48E-08	1.355

It is seen that the dry density of the soil increased from 18.3 kg/cm<sup>3</sup> to 19.21 kg/cm<sup>3</sup> and the permeability decreased from  $1.8 \times 10^{-4}$  to  $1.6 \times 10^{-4}$  cm/s. It is seen that while the density changes from 18.3 kg/cm<sup>3</sup> to 19.21 kg/cm<sup>3</sup>, the permeability coefficient decreases rapidly and falls more slowly from 19.21 kg/cm<sup>3</sup> to 19.7 kg/m<sup>3</sup>. This indicates that the soil in its natural state is loose and highly permeable.

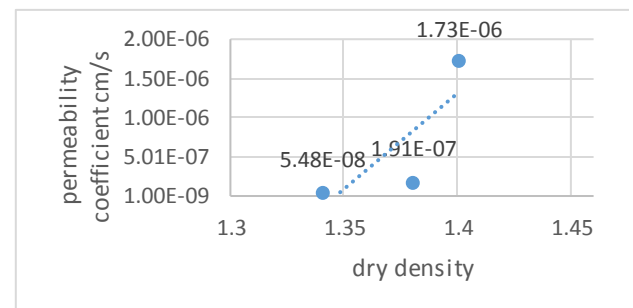


Figure III: coefficient of permeability and dry density (First Sample).

Table IX:  $k$  and  $\gamma_d$  for different ratios

Sample Ratio	Dry density (kg/cm <sup>3</sup> )	K (cm/s)
(95%S+5% B)	19.7	1.1
(%90S+10%B)	19.21	1.68
(85%S+15%B)	18.3	1.75

The graph of the old sample shows that the permeability coefficient increases when the dry unit weight increases.

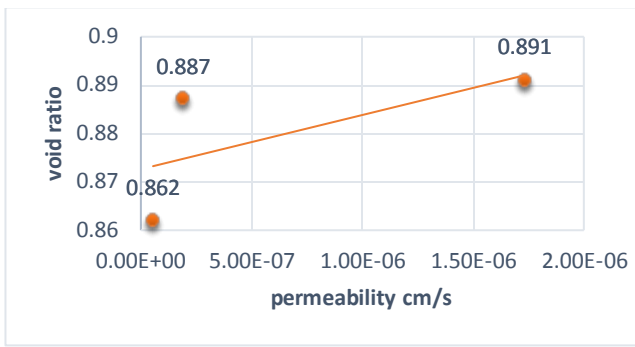


Figure IV: permeability and void ratio for old sample (composite).

Table X: Permeability and void ratio for different samples.

Ratio	K (cm/s)	Void Ratio
80%S+20%B	1.73E-06	0.891
70%S+30%B	1.91E-07	0.887
60%S+40%B	5.48E-08	0.862

The void ratio is increased owing to permeability, the larger quantity of voids in the particles, and the flow area in the samples. In bentonite, flow in already small channels is further hindered because some of the water in the voids is adsorbed or adsorbed on the bentonite particles, reducing the flow area and further restricting the flow. Therefore  $K_{\text{Bentonite}} \lll K_{\text{sand}}$ .

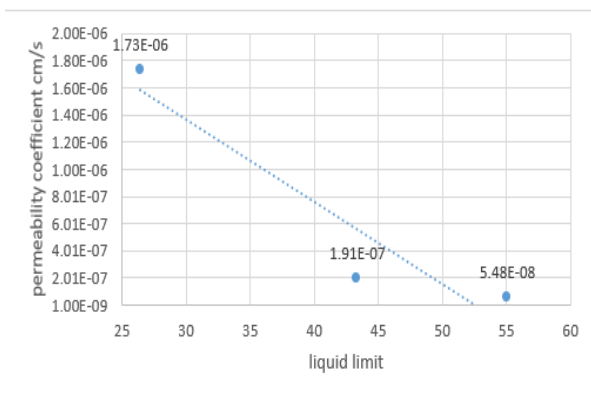


Figure V: liquid limit and permeability coefficient.

Table XI: k and LL values for samples.

Ratio	Permeability (cm/s)	LL
80%S+20%B	1.73E-06	26.44
70%S+30%B	1.91E-07	43.39
60%S+40%B	5.48E-08	55.1

The liquid limit will grow and the coefficient of permeability will decrease. The permeability coefficient also decreases with the increase of the plasticity index. Since the plastic limit approaches nearly zero when the sand content in the sample is high. The explanation for this behavior is that as the plasticity of the soil sample increases, the soil particles prefer to keep the water between the voids, decreasing the possibility of water passing through the voids. That's why plastic clayey soils are used as soil barriers in several geotechnical projects such as municipal landfills and the core of the dams to impede infiltration.

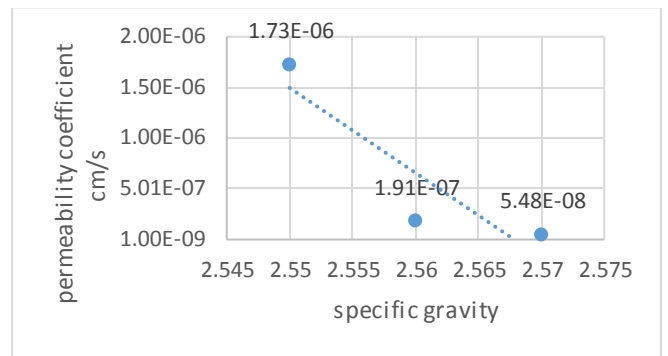


Figure VI: permeability coefficient and specific gravity.

Table XII: Gs and k for samples.

Ratio	Permeability coefficient (cm/s)	Specific gravity
80% S+20%B	1.73E-06	2.55
70% S+30%B	1.91E-07	2.56
60% S+40%B	5.48E-08	2.57

The increasing in specific gravity was observed with increasing the bentonite content due to high specific gravity of bentonite. As the specific gravity increases, the permeability coefficient will be decrease.

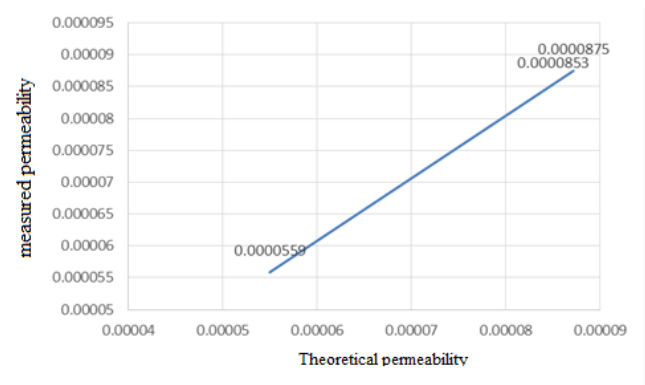


Figure VII: measured permeability and theoretical permeability (n).

Figure VII shows that in most cases the measured permeability values are greater than the theoretical values.

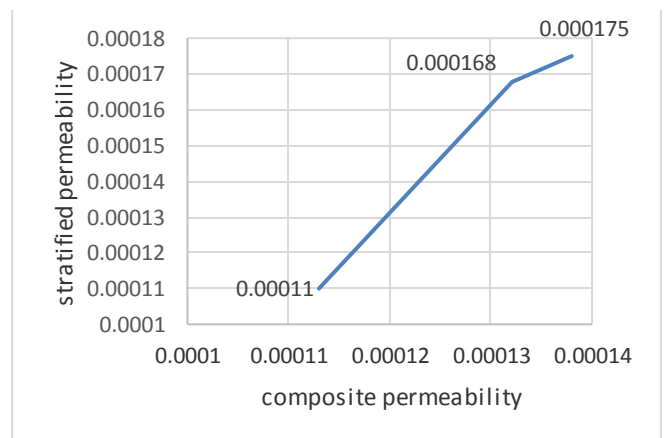


Figure VIII: Stratified and Composite permeability.

Table XIII: k and LL of samples.

Ratio	Composite Permeability (cm/s)	Stratified Permeability (cm/s)
95% S+5%B	1.38E-04	0.000175
90% S+10%B	1.32E-04	0.000168
85% S+15%B	1.13E-04	0.000113

Considering the results for layered soils reveals that the permeability decreases as the bentonite increases, and the permeability of the composite increases as the bentonite increases.

#### IV. Results

For bentonite and sandy soil, atterberg limit, specific gravity, sieve analysis, proctore test and laboratory study were carried out for two layers systems with different soil types, layer type, varying ratios, position done and the equivalent permeability coefficient was found to be different from the value calculated from Darcy's law. The exit layer's permeability checks if the measured permeability for a stratified layer is above or below the theoretical values. The permeability measured in this document is larger than that measured in theory. In the layered system, in addition to void ratio, thickness, and soil type, the permeability coefficient of a soil seems to be a function of the interaction between the soil and its surrounding soil(s).

And so it should be considered that the permeability coefficient of a soil in a layered system also depends on the direction of flow, its relative position, and the thickness of the layer. This work is purely experimental and opens up space for further study and therefore deriving a mathematical equation for layered soils.

- For stratified deposition, the permeability of the exit layer controls whether the measured permeability is greater or less than the theoretical permeability.
- It is seen that the permeability coefficient increases when the dry unit volume weight increases compared to both the old and new samples.
- As the coarse-grained material increases, the void ratio and the permeability increases too.
- The permeability coefficient decreases as liquid limit increases.
- Plasticity index and plastic limit increase with the permeability coefficient decreases.
- As the specific gravity increases, the permeability coefficient decrease.

- The measured permeability greater than theoretical permeability.
- For layered soils, the permeability decreases as the bentonite increases, and the permeability of the composite increases as the bentonite increases.
- The flow rate increases with the increase in the hydraulic gradient rate.

#### References

- [1] Kozłowski, T., & Ludynia, A. (2019). Permeability Coefficient of Low Permeable Soils as a Single-Variable Function of Soil Parameter, *Water* 11(12).
- [2] Stephenson, R., & Uppot, J. (January 1989). Permeability of Clays Under Organic Permeants. *Journal of Geotechnical Engineering*, 120-131.
- [3] haug, m. (February 1990). Evaluation of In Situ Permeability Testing Methods. *Journal of Geotechnical Engineering*, 116-297.
- [4] Prakash, K., & Sridharan, A. (2002). Permeability of two-layer soils. *Geotechnical Testing Journal*, 443-448.
- [5] Gustavo, F. S., & Elsharief, A. (July 15, 2004). Effects of Lime on Permeability and Compressibility of Two Tropical Residual Soils. *Journal of Environmental Engineering*.
- [6] Chegenizadeh, A., & Nikraz, H. (2011). Permeability test on reinforced clayey sand. *World Academy of Science, Engineering and Technology*, 130-133.
- [7] Sridharan, A., & Prakash, K. (2013). Permeability of Layered Soils: An Extended Study. *Geotechnical Testing Journal*, 31(5):1639-1644.
- [8] <https://www.geoengineer.org/education/laboratory-testing>.

# An AI-based Embodied Digital Human Assistant for Information in University

Munia AlKhalifa  
Department of Computer Engineering  
Karabuk University  
Karabuk, Turkey  
munia.ak@outlook.com

Kasım Özacar  
Department of Computer Engineering  
Karabuk University  
Karabuk, Turkey  
kasimozacar@karabuk.edu.tr

**Abstract**—Empowered by artificial intelligence (AI), digital assistants are taking an essential role in our lives as they are serving the needs of people within many domains such as providing customer service and translating. The two well-known types of assistants, which are text-based and voice-operated, are agents that answers the question or serve the requests of users depending on data that is given to them. Techniques used for building these agents differ depending on multiple measures and the methods these agents follow serve varies from being basic-simple rules to state-of-art techniques. However, to achieve a natural and accurate interaction like we do in our everyday life, we contribute with a voice-based digital assistant integrated into a virtual human whose aim is to serve at university working as a friendly assistant that answers the questions of students, newcomers, or visitors. We built and trained multiple neural networks models and combined them to have a human assistant responding any query related to the university while expressing and being interactive.

**Keywords**—Human-Computer Interaction, Artificial Intelligence, digital assistant, chatbot, deep learning

## I. Introduction

AI-based digital assistants have been wide spread as their affordability and efficiency make them a key element and a recognized example in all industries and empowered by simple to advanced artificial intelligence techniques. These agents are also known as digital assistants which can perform different tasks as well as are capable of mimicking human impressions in conversations along with providing the requested service in applications like ecommerce, information retrieving, and education [1].

A big challenge is the goal of developing assistants that to interact naturally with human and at the same time generate much more natural or human-like conversations making them indistinguishable from that of a human during a normal open-domain or closed-domain conversations that are used in providing service and help to users. In addition to conversation flow and service provision, another challenge is to make the chatbot acts not only as a tool, but also as a friend [2].

During the last two decades technologies such as speech recognition, natural language understanding, and text to speech synthesis have been the interest for much research resulting in well-known digital assistants like Apple's Siri [3]. Scenarios for digital assistants have been extended in a variety of areas such as tutoring [4], health care [5], forecasting [6], translation [7] and navigation [8]. Digital assistants can answer both simple and complex questions, provide information, recommendations, initiate conversation occasionally and make predictions.

Additionally, to the importance of questions responding, face to face communication achieves notable effects on people, thus, this will offer better performance and more convincing interactions. Therefore, in this paper we build a virtual human as a digital assistant which works a guide and information provider about our university by answering the questions of users and interacting with them.

The contributions of the paper are as follows: 1) we introduced our idea of a university digital assistant, 2) we go through discussing some previous research of same field 3) we explain the methodology of our work 4) we explain about tests and experiments 5) we discuss our work and describe the drawbacks 6) we conclude our paper.

## II. Related Work

Conversations and communication between human and machine have been receiving attention since the early 1960s. ELIZA, was the first natural language processing program and chatbot who worked as psychotherapist and used simple pattern matching [9], then came IBM's Shoebox-activated calculator and by the year's big companies started developing speech recognition systems and machines.

Another chatbot was launched on platforms like MSN Messenger, was assigned with simple tasks such as checking weather, converse with users, and looking up facts [10]. Parry [11] is an improvement over ELIZA having its own personality. In 2000, 2001, and 2004 a chatbot named ALICE won the Loebner prize due to its high similarity to humans [12] though it relies on simple pattern-matching algorithm

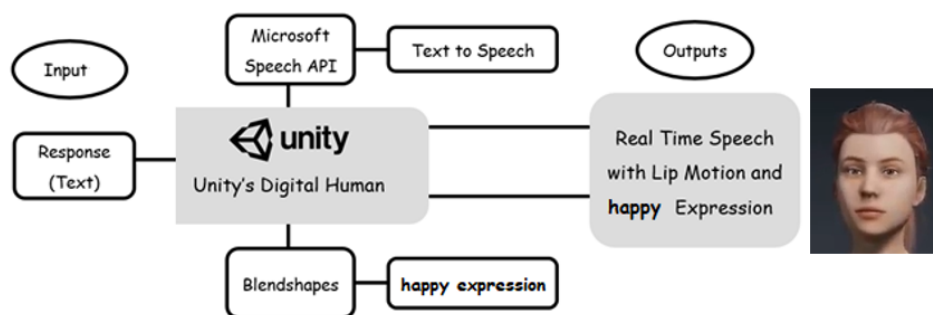


Figure I. (B). System workflow inside Unity.



based on the Artificial Intelligence Markup Language (AIML) [13].

Afterwards, many virtual personal assistants like IBM Watson [14], Amazon Alexa [15], Google Assistant [16], Microsoft Cortana [17], and Apple Siri [18] came to the life. By far we can see that chatbots, can be text-based or vocal and they are utilized into different industries including marketing, supporting systems, education, healthcare, cultural heritage, and entertainment. Therefore, our digital assistant who works a guide and answers questions while interacting with human.

### III. Methodology

We aim at a digital human assistant that answers any given question regarding the university. We summarize the architecture in Figure I (A) and (B). In the coming sections, we explain each part in the system.

#### A. Speech Recognition

Speech recognition or Speech to Text can be defined as converting the speech sound signal into instructions or words to give the machines the ability to respond to these commands [19, 20]. We deploy this technique in our work because the virtual assistance needs to understand what the user asks when speaking which leads to the need of converting the user's speech into text to apply the natural language understanding (NLU) techniques on the text.

Many APIs have been introduced for speech recognition such as DeepSpeech [21]. However, we use a very simple script to convert speech into text with the help of Google Speech Recognition API [22]. This script will be combined with the chatbot to understand the speech then respond as a text, then again converted back to speech.

#### B. Question Answering

Before building a chatbot, its objective must be decided to pick the correct type of chatbot that suits the business or task. Classifying the bot depends on varied parameters like response generation technique, the goal, input processing, the domain of knowledge, the provided service and finally the method chosen for building. [23]

We decided what kind of dialog system we will build for our agent starting with the knowledge domain, as we will be providing information related only to the university then that decides that domain will be closed-domain instead of the open-domain.

Then considering goal classification, we defined what goal the bot should achieve. Chatbots, like FAQ bots, can be

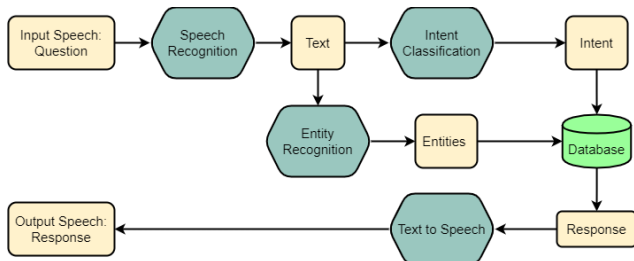


Figure I. (A). System workflow between the models.

informative built for providing information from a source, or could be chat based which talks to the user typically as a human by responding with the correct sentence, or can be task-based like bots helping in booking a flight or like a FAQ chatbot. Another measure for deciding the bot type is the input processing and response generation. From the three response generation models: rule-based, retrieval-based, and generative model [24], we picked the retrieval-based model.

After specifying the correct type, we started preparing our own dataset of questions and database of responses. The bot will retrieve the response from the response candidates in the database to answer the user's question considered as queries. For achieving this purpose, we will explain each part included in the bot. The four essential concepts are: intents, entities, Named Entity Recognition (NER), Intent Classification. We display in figure II a conversation example held between a user and the bot in the as textual interface before combining with the assistant virtually in order to check how well the chatbot is answering.

We created our own dataset that is a collection of different questions related to Karabuk University that reached 300 questions in total. We made two copies of the dataset, one that suites the first model and the other suits the second model but both having same questions. Relatively we created the database of answers to these questions. For our trainings we divide the dataset as 80% for training and 20% for validation.

#### C. Intent Classification

Table I. Intents and Entities.

Intents	Entities
AboutErasmus	institution
AboutFaculties	faculty
AboutLocation	major
AboutMajors	action
AboutResidence	money
AboutTransportation	service
AboutUniversity	residence
Available Activity	place
Evaluation	location
FacultyMajor	activity
Fees	person
Greet	program
IsMajorAvailable	ranking
Staff	city
WhyKarabuk	

The intent classification is an NLP problem and more specifically NLU problem which aims to understand the intention of the user through machine learning model that will recognize what the user originally demands and consequently decide what step should be taken next. The intent classification is very essential part in this system and is considered as sentence classification that one label or multiple labels will be detected from the sentences.

[Assistant]: Hello and Welcome to Karabuk University, I'm your assistant and friend. You can ask me anything you would like to know about the university and I will be gladly helping you.

Enter Your Message: Hello, I'm a new student and I would like to know generally about karabuk university.

[Assistant]: Karabük University is a public higher education institution in Turkey. KBU opened its doors in 2007. Inside the city of Karabuk the university has built its own suburban campus. Karabük University never leaves the national rankings of the top universities of Turkey.

Enter Your Message: nice.. so do you provide engineering majors?

[Assistant]: of course we do. We have computer engineering, biomedical engineering, electrical engineering, environmental engineering, civil engineering, mechanical engineering, mechatronics engineering, metallurgical and materials engineering and transportation engineering.

Enter Your Message: I want to know how to apply for example what is the process and the cost of tuition?

[Assistant]: The admission is based on the results of the examinations. In order to get accepted the applicant just have to match the criteria set by the university. The academic year in KBU is divided to 2 academic semesters. KBU has affordable bachelor's programs that cost less than 2,500 USD/year. A master's degree of KBU will cost a student a significant sum of 2,500 USD per year. Is among the educational institutions that allow the students to enroll in online programs. On the official website (<http://www.karabuk.edu.tr>) of the university you can find a more accurate information regarding the tuition fees of different programs.

Enter Your Message: thanks for the information Goodbye.

[Assistant]: take care and have a nice day!

Figure II. Conversation example between user and the chatbot.

Typically, many techniques are suggested for intent recognition like the RNN [25] and SVM [26]. However, in our work we build a Convolutional Neural Network (CNN) for sentence classification [27]. Following up with the original paper, we build a convolutional neural network on top of word embeddings. To be able to solve text classification problem with CNN we need to use pre-trained word embeddings. Word embedding is NLP techniques for mapping semantic context into geometric space known as embedding space. Each word in a dictionary will be assigned with a numeric vector where the distance between each two vectors represents a part of the relationship semantically between the two assigned words. For example school and book can be related words thus they will be embedded closed to each other, but apple and cat are not related semantically thus they will be represented as distant word vectors in the embedding space.

Computation of word embedding is done through neural networks such as the word2vec technique which is used in this model. A lot of word embeddings datasets can be downloaded, for our work we will be using the 100-dimensional GloVe embeddings (Global Vectors for Word Representation) of 400k words computed on of English of Wikipedia.

To build the intent classification model we start with converting our dataset text samples into word indices then

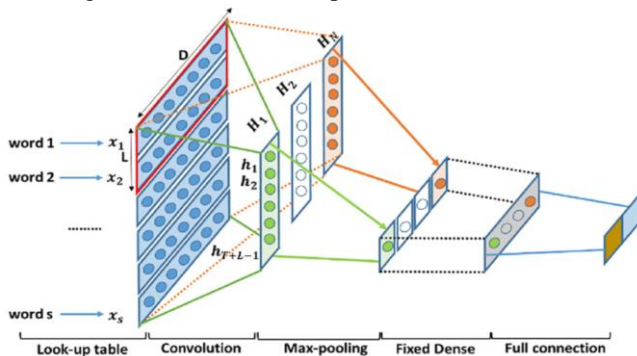


Figure III. Standard CNN on text classification.

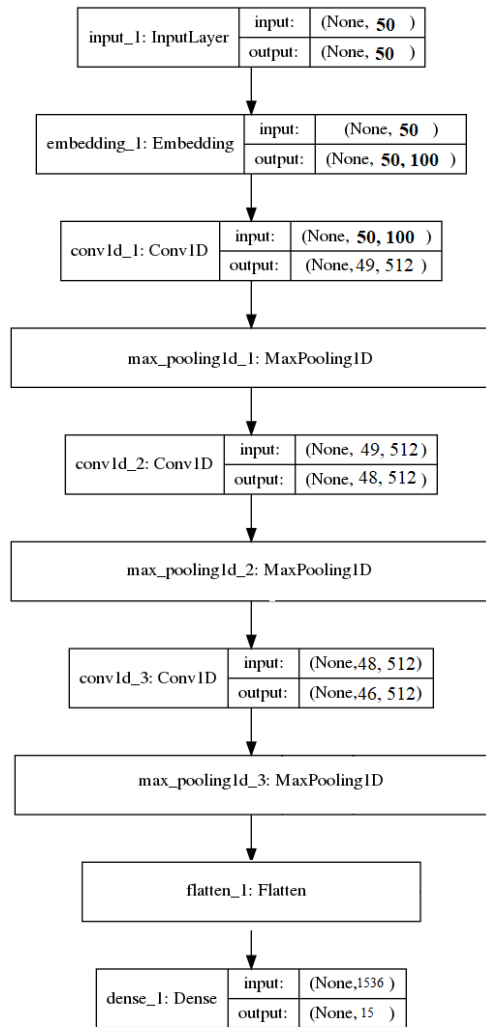


Figure IV. CNN layers for text classification.

preparing embedding matrix that has the embedding vector at index  $z$  for the word of index  $z$  in the word index. We load the matrix to Embedding layer of Keras and build on top of it a simple 1D convolutional neural network.

The basic architecture of CNN for text classification is shown in figure III. The CNN in our work is displayed in figure IV, as described by [27] it consists of three 1D convolutions and one dense layer. We train the model using the following hyperparameters:

- 50 epochs
- filter\_sizes = [2, 3, 5]
- number of filters = 512
- optimizer = Adam
- loss function = categorical\_crossentropy
- ReLU and softmax activation functions

Output will be classified into one of the 15 intents shown in Table I. The model achieved 99% accuracy. We displayed the evaluation metrics of the model in Table II.

## D. Named Entity Recognition

Named-entity recognition is a Natural Language Understanding (NLU) problem and means information extraction but for finding categories known as entities within the text [28]. Entity is a word considered as a parameter value that is extracted from context. While intents refer to user's main goal, the entities work as keywords that refer to meaningful or important things and used by the users to describe what they want, or in other words, to describe their intents. Entities can be system-defined like data references [29] but for our problem we define our own entities. General entities examples are person, location, organization, city, date, etc. depending on the industry, and they are called as domain entities which gets tagged from an input sentence [30].

In our work we got 14 entities in total and are displayed in the Table I. For entity extraction many libraries and frameworks have been introduced like Snips [31] and Rasa [32]. We used an NLP open-source library called SpaCy [33] that is alternative to the popular NLTK [34] and includes pretrained machine learning models.

SpaCy made NLP easier in Python by providing new pipelines based on transformer which improved the accuracy, efficiency and adaptability of SpaCy especially in the third version. The NER model of SpaCy assigns label to contiguous tokens groups. The NER model of SpaCy consists of the following:

- wordly-wise word embedding technique through subword features and Bloom embedding
- deep convolutional neural network with residual layers
- named entity parsing with modern transition dependent approach

We labeled the entities in the question dataset using SpaCy NER annotator [35] and an example in Figure V shows how the entities look like in a question. The NER model gets the sentence and extracts keywords that belong to proper domain entities and using them along with the intent a response will be retrieved from the database.

For fine tuning with SpaCy's NER we start with creating the json file of our questions dataset with defined entities. Afterwards we convert it to SpaCy file and create the configuration and using SpaCy's built in config files and this is simply provided with one command line. Once done we begin training with a SpaCy pipeline imported from their library. The models was trained for 100 epochs and was evaluated on 20% of dataset and new data prediction. We show the evaluation metrics of the model in Table II.

## E. Combining with Unity

To combine the models with the virtual human project [36] we connect them to Unity game engine using UDP communication. In this procedure we send the data from python model by sockets to Unity that has UDP client that reads received data from the socket. We summarize this in Figure VI. Python server will send the response as text to Unity's C# client in real time. In unity's virtual human we generate the speech, lip motion and smiling facial expression when interacting with user.

## IV. Preliminary Study

For our experiments we conducted a preliminary test applying qualitative and quantitative analysis with 12 users in total between the ages of 23 and 45 (9 of them from our university). For the qualitative analysis we have interviewed 5 users (1 female) by asking them three questions and then we built high-level themes from their answers that will be used in the second quantitative test. In the quantitative analysis we conducted a survey with 7 users (2 females) after letting them test the system by interacting with the assistant, and in the survey they were asked to give feedback over statements by rating them. In both of the study steps all participants were either staff, or instructors in the computer engineering department, or lecturers from other departments.

### A. Qualitative Test

For the interviews, we applied them in different rooms of the university and during them we concentrated on providing an undisturbed surrounding where no other persons than the interviewer and the participant were present. We interviewed

I would like to know about **computer engineering** in **Karabuk University**.

Entity: Major      Entity: Institution

Figure V Entities extracted from a sentence.

the participants individually by asking them 3 questions in sequence. First questions is "how advantageous do you see the 3D virtual human assistant? And why?" second question is "what kind of improvements should be added to any virtual assistant?" and the third question was "what are the most common problems that may occur during the communication with a machine" They were permitted to answer with whatever comes to their minds and then it will be edited after the interview. The interviewer collected their answers on Microsoft Word. We aimed for deriving high-level themes

Table II. Evaluation Metrics on neural network models.

	F1	Precision	Recall
Intent Classifier	99	99	99
Entity Recognizer	99	99	98

from the participants' responses which will be used in the survey part.

### B. Quantitative Test

For this part we held a two steps test with the second group of participants and as same as the previous group, they were conducting the tests in undisturbed rooms. For this part we firstly let the participants individually try the system by running the virtual human assistant with the use of virtual reality (VR) technology, and then we apply an online survey consisting of 5 questions in total, three of which were based on the topics introduced previously in the qualitative test analysis.

We conducted the survey on SurveyMonkey and evaluated the statements on the Likert scales that includes "strongly disagree," "disagree," "neutral," "agree," and "strongly agree." We display the results of the survey for each statement separately in the percentages form.

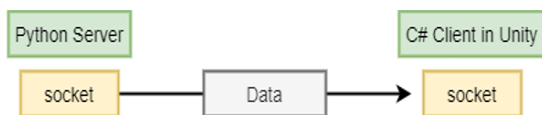


Figure VI. UDP connection between Python and Unity's C#

Our goal of using these evaluations is to quantify the range of participants' agreement with the statements built on the interviews part by providing strong evidence rather than simply applying normal feedback survey.

### C. Results

For the interview responses qualitative analysis, we built the themes depending on keywords the frequently occurred in the answers namely fast /quickly/quick, information/data, human/human-like/humane/humanoid, interaction/interactive/interact, intent/intention, recognize/recognition and many more. Depending on the count of words we constructed the themes. These themes are the following:

- Theme 1, human-like interaction. (Qualitative based)
- Theme 2, quick responses. (Qualitative based)
- Theme 3, persuasive in interaction. (Qualitative based)
- Theme 4, fast recognition of intent.
- Theme 5, highly informative and clear.

We summarize the results in figure VII through a bar chart that shows the number of participants who chose a feedback on every theme. For instance 3 of users strongly agree on theme 4, while 3 others agree and only on feels neutral.

### V. Discussion

Introducing an informative digital assistant was the key insights in this research that will serve students, visitors, or newcomers to the university by answering any question related to our university. To achieve this, we integrate several models to a 3D virtual character. These models are speech to text model, question answering model, and sound synthesizing. We believe that building such assistant will benefit the school by attracting the attention of users and provide all needed information to them as a normal human. Additionally, the assistant can help each user individually.

However, we still believe that the assistant should provide more changeable responses and ask back the user in more flexible chat flow till understanding the intention of user. Therefore, future work will be enhancing the dataset, the conversation flow, provide more interactive gestures and advanced service by the agent such as providing a tour to the school visitors.

### VI. Conclusion

Digital assistants are being dominant applications in real life that serve the community easily using artificial intelligent techniques. Additionally, virtual humans are playing a very important role in the human-computer interaction allowing people to interact with the system. In this paper our purpose was to build an assistant that helps in university queries. We combined multiple models each one achieving task such as a model for speech recognition, a model for entity recognition and a model for intent classification. Moreover, we deploy voice synthesizing plugin in unity with these models. The

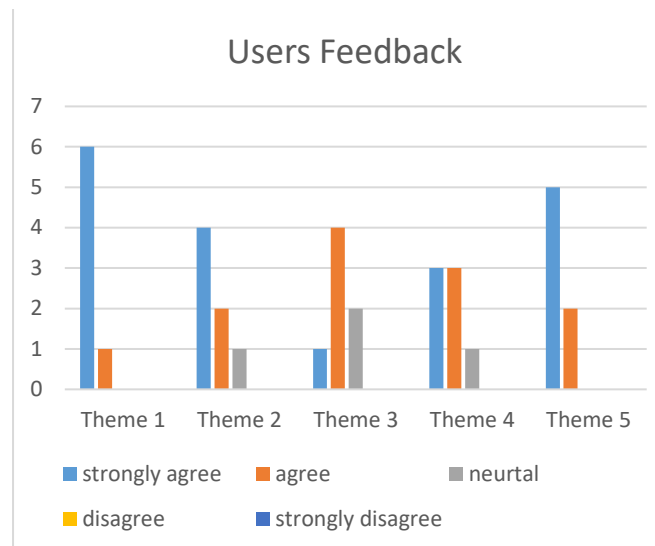


Figure VII. Diagram shows the number of users who strongly agree, agree, disagree, strongly disagree and had neutral feedback on each theme.

reason for choosing a virtual human to be a digital assistant is because of the importance that expressions and face-to-face communication play in a conversation. Therefore, this project is a contribution for serving at university but built with the state-of-art techniques.

Our future plan to add more expressions, gestures and increasing the size of the dataset to improve the project.

### References

- [1] Abu Shawar, B.A., Atwell, E.S.: Chatbots: are they really useful? *J. Lang. Technol. Comput. Linguist.* 22, 29–49 (2007)
- [2] Brandtzaeg, P.B., Følstad, A.: Why people use chatbots. In: Kompatsiaris, I., et al. (eds.) *Internet Science*, pp. 377–392. Springer, Cham (2017). [https://doi.org/10.1007/978-3-319-70284-1\\_30](https://doi.org/10.1007/978-3-319-70284-1_30)
- [3] Siri. <https://www.apple.com/siri/>
- [4] Braun and N. Rummel, "Facilitating Learning From Computer-Supported Collaborative Inquiry: the Challenge of Directing Learners' Interactions To Useful Ends," *Research and Practice in Technology Enhanced Learning*, vol. 05, no. 03, pp. 205, 2010.
- [5] D. Coyle, G. Doherty, M. Matthews, and J. Sharry, "Computers in talk-based mental health interventions," *Interacting with Computers*, vol. 19, no. 4, pp. 545–562, 2007.
- [6] Zue, S. Seneff, J. R. Glass, J. Polifroni, C. Pao, T. J. Hazen, and L. Hetherington, "J UPITER : A TelephoneBased Conversational Interface for Weather Information," *IEEE Transactions on Speech and Audio Processing*, vol. 8, no. 1, pp. 85–96, 2000.
- [7] M. Kolss, D. Bernreuther, M. Paulik, S. Stucker, S. Vogel, and A. Waibel, "Open Domain Speech Recognition & Translation: Lectures and Speeches," in *Proceedings of ICASSP*, 2006.
- [8] R. Belvin, R. Burns, and C. Hein, "Development of the HRL route navigation dialogue system," in *Proceedings of ACL-HLT*, 2001, pp. 1–5.
- [9] Weizenbaum, J.: ELIZA—a computer program for the study of natural language communication between man and machine. *Commun. ACM* 9, 36–45 (1966). <https://doi.org/10.1145/365153.365168>
- [10] Adamopoulou E, Moussiades L. An Overview of Chatbot Technology. *Artificial Intelligence Applications and*

- Innovations. 2020;584:373-383. Published 2020 May 6. doi:10.1007/978-3-030-49186-4\_31
- [11] Colby, K.M., Weber, S., Hilf, F.D.: Artificial paranoia. *Artif. Intell.* 2, 1–25 (1971). [https://doi.org/10.1016/0004-3702\(71\)90002-6](https://doi.org/10.1016/0004-3702(71)90002-6)
- [12] Wallace, R.S.: The anatomy of A.L.I.C.E. In: Epstein, R., Roberts, G., Beber, G. (eds.) *Parsing the Turing Test: Philosophical and Methodological Issues in the Quest for the Thinking Computer*, pp. 181–210. Springer, Cham (2009). [https://doi.org/10.1007/978-1-4020-6710-5\\_13](https://doi.org/10.1007/978-1-4020-6710-5_13)
- [13] Marietto, M., et al.: Artificial intelligence markup language: a brief tutorial. *Int. J. Comput. Sci. Eng. Surv.* 4 (2013). <https://doi.org/10.5121/ijcses.2013.4301>
- [14] IBM Watson. <https://www.ibm.com/watson>
- [15] What exactly is Alexa? Where does she come from? And how does she work? <https://www.digitaltrends.com/home/what-is-amazons-alexa-and-what-can-it-do/>
- [16] Google Assistant, your own personal Google. <https://assistant.google.com/>
- [17] Personal Digital Assistant - Cortana Home Assistant – Microsoft. <https://www.microsoft.com/en-us/Cortana>
- [18] Siri. <https://www.apple.com/siri/>
- [19] Hui Liu, Chapter 1 - Introduction, Editor(s): Hui Liu, *Robot Systems for Rail Transit Applications*, Elsevier, 2020, Pages 1-36.
- [20] Zwass, Vladimir. "Speech recognition". *Encyclopedia Britannica*, 10 Feb. 2016, <https://www.britannica.com/technology/speech-recognition>. Accessed 12 June 2021.
- [21] Hannun, C. Case, J. Casper, B. Catanzaro, G. Diamos, E. Elsen, R. Prenger, S. Satheesh, S. Sengupta, A. Coates, et al. *Deep speech: Scaling up end-to-end speech recognition*. arXiv preprint arXiv:1412.5567, 2014
- [22] <https://pypi.python.org/pypi/SpeechRecognition/>
- [23] Adamopoulou, Eleni & Moussiades, Lefteris. (2020). An Overview of Chatbot Technology. 373-383. 10.1007/978-3-030-49186-4\_31.
- [24] Hien, H.T., Cuong, P.-N., Nam, L.N.H., Nhung, H.L.T.K., Thang, L.D.: Intelligent assistants in higher-education environments: the FIT-EBot, a chatbot for administrative and learning support. In: *Proceedings of the Ninth International Symposium on Information and Communication Technology*, pp. 69–76. ACM, New York (2018)
- [25] Abiodun, Oludare Isaac; Jantan, Aman; Omolara, Abiodun Esther; Dada, Kemi Victoria; Mohamed, Nachaat Abdelatif; Arshad, Humaira (2018-11-01). "State-of-the-art in artificial neural network applications: A survey". *Heliyon*. 4 (11)
- [26] Cortes, C., Vapnik, V. Support-vector networks. *Mach Learn* 20, 273–297 (1995). <https://doi.org/10.1007/BF00994018>
- [27] Kim, Yoon. (2014). Convolutional Neural Networks for Sentence Classification. *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing*. 10.3115/v1/D14-1181.
- [28] Perera N, Dehmer M, Emmert-Streib F. Named Entity Recognition and Relation Detection for Biomedical Information Extraction. *Front Cell Dev Biol.* 2020 Aug 28;8:673. doi: 10.3389/fcell.2020.00673. PMID: 32984300; PMCID: PMC7485218.
- [29] Ramesh, K., Ravishankaran, S., Joshi, A., Chandrasekaran, K.: A survey of design techniques for conversational agents. In: Kaushik, S., Gupta, D., Kharb, L., Chahal, D. (eds.) *ICICCT 2017*. CCIS, vol. 750, pp. 336–350. Springer, Singapore (2017). [https://doi.org/10.1007/978-981-10-6544-6\\_31](https://doi.org/10.1007/978-981-10-6544-6_31)
- [30] Jung, S.: Semantic vector learning for natural language understanding. *Comput. Speech Lang.* 56, 130–145 (2019). <https://doi.org/10.1016/j.csl.2018.12.008>
- [31] <https://github.com/snipsco/snips-nlu>
- [32] Bocklisch, Tom & Faulker, Joey & Pawlowski, Nick & Nichol, Alan. (2017). *Rasa: Open Source Language Understanding and Dialogue Management*.
- [33] <https://github.com/explosion/spaCy>
- [34] Bird, S., Klein, E., & Loper, E. (2009). *Natural language processing with Python: analyzing text with the natural language toolkit*. "O&#x27;Reilly Media, Inc."
- [35] <https://github.com/ieriii/spacy-annotator>
- [36] <https://github.com/GeoffreyGorisse/VHProject>

# Visual Question Answering for Medical Image Analysis based on Transformers

Hanan Othman  
Department of Computer Engineering  
King Saud University  
Riyadh, Saudi Arabia  
439203835@student.ksu.edu.sa

Yakoub Bazi  
Department of Computer Engineering  
King Saud University  
Riyadh, Saudi Arabia  
ybazi@ksu.edu.sa

Mohamad Alrahal  
Department of Applied Computer Science  
King Saud University  
Riyadh, Saudi Arabia  
mmalrahal@ksu.edu.sa

**Abstract**— Health care has been revolutionized over the past decades in conjunction with new discoveries and technological advancements. One of those areas that have been rapidly evolved is medical imaging that plays a significant role in screening, early diagnosis, and treatment selection. Artificial Intelligence (AI) has been utilized to support the physician's decisions related to medical imaging. Recently, medical visual question answering (VQA) has been utilized to predict the right answer for a given medical image accompanied with a clinically relevant question to support the clinical decision. However, the validity of medical VQA is still not proven. In this paper, we proposed a full transformers architecture for generating answers given the question and image. We extracted image features using the data-efficient image transformer (DeiT) model and bidirectional encoder representations from the transformers (BERT) model for extracted textual features. We also applied a concatenation to integrate the visual and language features. The fused features were then fed to the decoder to predict the answer. This model established new state-of-the-art results 61.2 in accuracy and 21.3 in BLEU score in the PathVQA data set.

**Keywords**—Transformers, visual question answer, medical image, vision transformers

## I. Introduction

Health care has been revolutionized over the past decades in conjunction with new discoveries and the advancement of technology. As a result, the physician has been started to adapt new modalities to optimize patient care. One of those updated modalities is using medical images. Thus, medical images play a significant role in screening, early diagnosis, and during surgeries such as cardiac catheterization [1]. Over recent years, artificial intelligence (AI) has started incorporating into the medical field and provided various innovation initiatives. Thus, AI has been utilized in a health care setting, such as advancement in diagnosis, treatment personalization, and electronic health recording. Many studies have indicated that the integration of AI in medical diagnosis programs had increased the accuracy, speed, and consistency of the diagnosis enhanced the prediction of patient outcomes and captured additional information as missed by the doctors [2], [3]. In this regard, several models have been proposed to help patients understand their physical conditions through visual inspection, such as image caption [4], image retrieval [5], visual question answering (VQA) [6], and visual question generation (VQG) [7].

Visual question answering (VQA) is a task that takes the medical image and clinical question about the image as input and produces natural language answers as output. This process shows great potential in providing medical assistance, such as helping patients get prompt feedback on their inquiries, making more informed decisions, supporting

the suitable utilization of medical resources, providing a second opinion to physicians in diagnosis, and reducing the high cost of training medical professionals.

In the literature, researchers have utilized different methods and models for medical VQA. Yan et al. [8] used BERT [9] and VGG-16 [10] with global average pooling (GAP)[11] strategy to extract question and image features, respectively. Then this is followed by co-attention to combine these features and a decoder to predict the answers. Chen et al. [12] used BioBERT for the questions and ResNet34 [13] to extract image features in which a bilateral-branch network (BBN) with a cumulative learning strategy [14] was used to fuse these features. Ren et al. [15] propose a model called CGMVQA that uses a multi-modal transformer architecture. Additionally, Khare et al. [16] used a similar architecture of CGMVQA with masked language modeling (MLM) and different datasets. Vu et al. [17] utilize a method denoted by Question-Centric Multimodal Low-rank Bilinear (QCMLB) that combines image and question features by applying high involvement to the query questions meaning.

However, VQA in the medical domain is still in its embryonic stage since the accuracy of previous methods has significantly lower than doctor's assessments, owing to the difficulty of answer evaluation and variety of answer expression. Thus, there is still a need to develop innovative techniques to overcome some of those limitations.

Recently, there is a growing body of literature that supports the utilization of Transformers in VQA. Thus, transformers have been initially used in natural language processing (NLP) tasks [18]. It is based entirely on attention mechanisms that use the encoder-decoder architecture. This attention focuses on specific parts of the input to get more efficient results. The main motivation of transformer models is to ability a long-range interaction between different sequence elements, unlike RNN. This motivation inspired Dosovitskiy et al. [19] to propose a convolution-free transformer called vision transformers (ViT) that tries directly to images by splitting the image into patches. These patches have been treated as tokens in NLP applications. These models led to very competitive results on the large dataset using extensive computing resources [20]. However, when ViT has trained in a small dataset, the model will not discover the properties of the image. Therefore, Touvron et al. [21] propose a new technique, called a data-efficient image transformer (DeiT), That requires less data (e.g., ImageNet1K) and less computing resources to produce a high-performance model. DeiT has the same architecture as the ViT model with knowledge distillation [22]. Knowledge distillation is a learning framework that uses student-teacher techniques by applying data augmentation coupled with

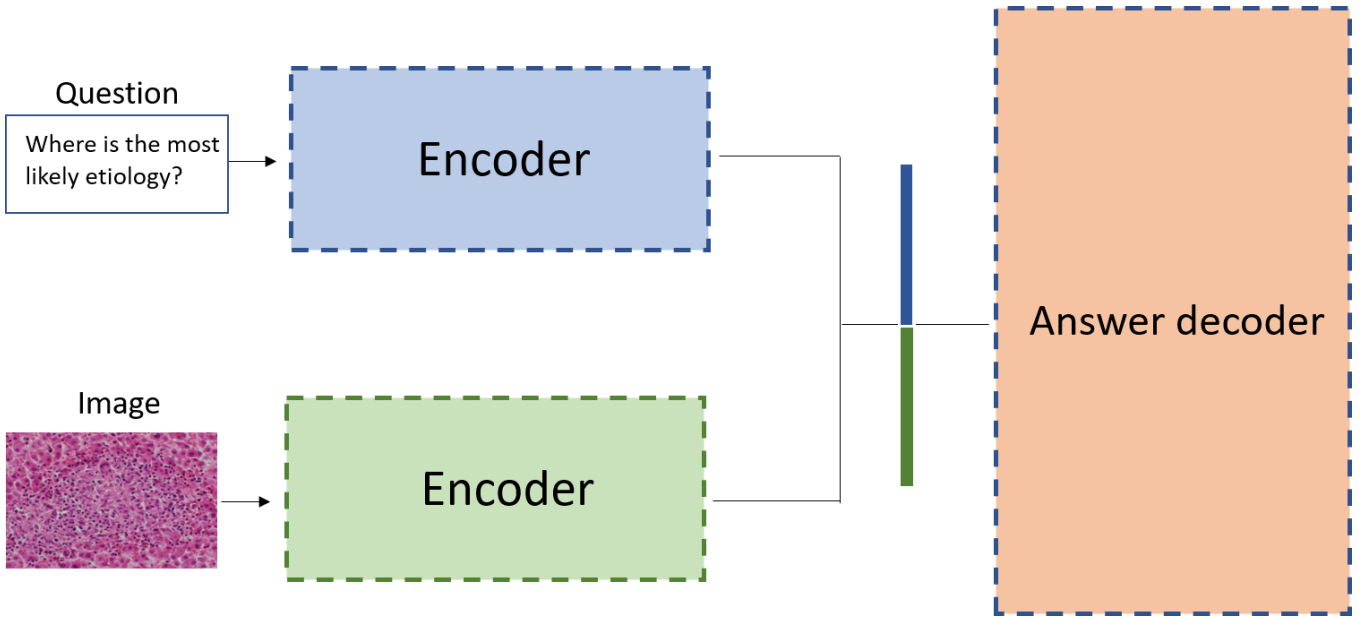


Figure I. An overview of our model that contain image encoder to encode the visual inputs, a text encoder to encode the language inputs, and a joint decoder to generate that answer

optimization tricks and regularization. This training strategy has shown excellent results on ViT for smaller datasets, particularly when the knowledge has been distilled from a convolutional neural network (CNN) teacher model [23].

Although previous studies used transformer as a single method, studies reported strong performance of multi-model learning such as VQA. Tan et al. [24] utilize a method denoted by learning cross-modality encoder representations from Transformers (LXMERT) that used a two-stream model with co-attention and only pre-trained the model with in-domain data. Lu et al. [25] used vision-and-language BERT (ViLBERT) with the same architecture with more complex co-attention, pre-trained with the out of domain data. Hu et al. [26] proposed a unified transformer (UniT) based on encodes each input modality with an encoder and a joint decoder that makes predictions for final outputs.

This paper proposes a full transformers encoder-decoder architecture for the VQA model in the medical domain. The encoder modules encode each input modality, and the decoder generates the answer word by word. The Image and question features have been extracted by using the DeiT model and BERT model, respectively. The extracted features for images and questions were fused using a fusion mechanism. Compared to previous work on multi-model learning with transformers, our work is the first one that trains on medical images.

The remainder of the paper is organized as follows: Section II describes the main methods based on transformers. In section III, we present the experimental results on the PathVQA dataset. Then we finally conclude and show future directions in Section IV.

## II. Proposed method

The proposed medical VQA framework is shown in Figure I, which consists of four parts: question encoding for extracting textual features of the given question, image encoding for capturing visual features of the given medical image, concatenation has used to fuse the visual and textual feature vectors to generate a jointed representation and

decoder for answer prediction. Detailed descriptions of the method have provided in the following subsections.

### A. Question encoder

We encode the question by using bidirectional encoder representations from Transformers (BERT). BERT is a transformer encoder model pre-trained on large corpora with masked language modeling and following sentence prediction tasks. It contains several blocks of Multi-Head Attention (MHA) followed by Feed Forward Neural Network (FNN).

Given the input questions  $Q_i$ , we tokenize it by adding a special classification token [CLS] followed by the WordPiece tokens and the separator token [SEP]. The token sequence is then added to word embedding to convert the tokens into vectors of dimension  $d_{model}$ . A positional embedding is added to each token to indicate its position in the sequence. Then the result will be fed to a transformer with several blocks. Each transformer block is composed of a multi-head self-attention (MSA), FFN, and normalization layer. The MSA block uses the self-attention mechanism to drive long-range dependencies between different words in the given text. Equation (1) Shows the details of the computations in one self-attention head (SA). First, the input sequence is transformed into three different matrices which are key vector  $K$ , query vector  $Q$ , and value vector  $V$  using three linear layers  $K \in \mathbb{R}^{d_{model} \times d_K}$ ,  $Q \in \mathbb{R}^{d_{model} \times d_Q}$ , and  $V \in \mathbb{R}^{d_{model} \times d_V}$  For  $i = 1, 2, \dots, h$  where  $h$  is the number of heads. The attention map was computed by matching the query matrix against the key matrix using the scaled-dot-product. The output is scaled by the dimension of the key  $d_K$  and then transformed into probabilities by a SoftMax layer. Finally, the result is multiplied with the value  $V$  to get a filtered value matrix which assigns high focus to more important elements.

$$Attention(Q, K, V) = softmax\left(\frac{QK^T}{\sqrt{d_K}}\right) \cdot V \quad (1)$$

Followed by FFN, which consists of two fully connected layers with a ReLU activation function in between. It can be formulated as:

$$FFN(x) = \max(0, xW_1 + b_1)W_2 + b_2 \quad (2)$$

In each sublayer, residual connections are existing, followed by layer normalization (LN) as described in the following equations:

$$x_{out} = \text{LayerNorm}(x_{in} + \text{Sublayer}(x_{in})) \quad (3)$$

where  $x_{out}$ ,  $x_{in}$  are the output and input of one sublayer, respectively, and the sublayer can be the attention layer or FFN.

### B. Image encoder

For image feature extraction, we encode the input image  $M$  with the DeiT model. The input image is divided into a sequence of patches with a dimension of  $H \times W \times 3$  (with three color channels). Each patch  $X$  has the dimension of  $(H \times W)/p$  Where  $p$  is the patch size ( $p=16$  in our experiment setting), these patches are comparable to word tokens in the original transformer. After that, we flatten each patch into embeddings by feeding them into a linear embedding layer  $E$  to map their dimension to the model dimension  $d_{model}$  and add position embedding to avoid the loss of the arrangement of the patches as in the original image. The resultant position-aware embeddings have appended with a learnable class token  $x_{class}$  and distillation token  $x_{distil}$  to the patch embeddings (Equation (4)). The two tokens and the patches embeddings interact with each other via the self-attention mechanism.

$$z_0 = [x_{class}; x_{distil}; x_p^1 E; x_p^2 E; \dots; x_p^m E] + E_{pos} \quad (4)$$

$$E \in \mathbb{R}^{(p^2 \cdot c) \times d_{model}}, E_{pos} \in \mathbb{R}^{(m+2) \times d_{model}}$$

The objective of the distillation token  $x_{distil}$  is to reproduce the hard label produced by the teacher network that allows the model to learn from the teacher's output while remaining complementary to the class token tasked to reproduce the actual label.

### C. Answer decoder

We generate the answer word by word for both the addition results and encoder output features as the input on the decoder side. The decoder consists of  $N$  layer, each layer containing masked multi-head attention followed by multi-head cross attention and FNN sequentially.

The masked multi-head attention block computes the attention vectors for current and prior words to avoid an information leak. An example shown in Figure II, in addition to the features of the image and the question, we use the features of all masked tokens to predict the first-word piece  $h_{a0}$  ("hydro"), then we use "hydro" and other following masked tokens to predict  $h_{a1}$ . Loop this process until the special token "[SEP]" is obtained.

## III. Experimental results

In this section, we present the dataset and evaluation metrics that used in the experiment. Then, the effectiveness of the proposed VQA in the medical image was analyzed. Finally, the preliminary results of this study were comparisons with the state-of-the-art model.

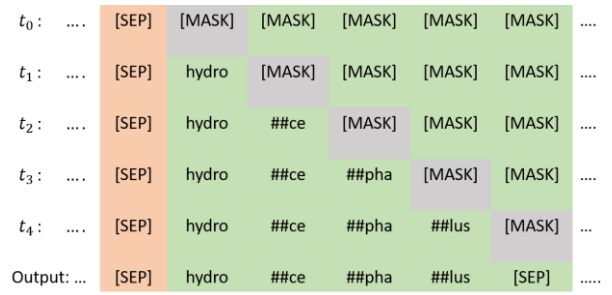


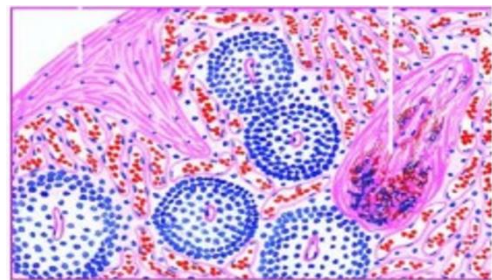
Figure II. The prediction process of generating the answer. Each line is the output of the previous time and the masked input of the current time, and the green box is the position of the predicted word piece

### A. PathVQA dataset

Compared with general domain images, medical images for VQA methods have a smaller amount of data and few datasets. We used the PathVQA dataset [27] that containing medical images accompanied by clinically relevant questions. This dataset has a total of 4,998 pathology images with 32,799 question-answer pairs and is split into three sets: training, test, and validation sets. The validation set has 987 images, the test sets contain 990 images, and the training set includes 3,021 images. The training set consists of 19,755 question-answer pairs, the validation set 6,279, and the test set consists of 6,761 question-answer pairs. Table I summarizes these statistics. As the example shown in Figure III. There are various questions related to multiple aspects of the image, such as location, shape, color, appearance, etc. The questions are classified into two types: open-ended questions with several varieties such as why, where, what, how, , etc., and close-ended "yes/no" questions. The task of the VQA-Med is to predict the most likely answer given the medical images.

Table I. Statistics of PathVQA data split

	Training set	Validation set	Test set
#Images	3,021	987	990
#Question-answer pairs	19,755	6,279	6,761



- Q1: **What** are dilated and congested?  
 Q2: **Are** the sinuses dilated and congested?  
 Q3: **Where** is increased fibrosis?  
 Q4: **Is** gamma-gandy body also seen?

Figure III. Exemplar of image with a various questions related to multiple aspects of the image



## B. Evaluation metrics

Accuracy [28] and BiLingual Evaluation Understudy (BLEU) [29] are commonly used as evaluation metrics in the VQA task. Accuracy measures the ratio between correctly predicted observations to total observations. Meanwhile, BLEU measures the similarity of predicted answers and ground-truth by matching n-grams.

## C. Results

Our model is trained using Adam optimizer with an initial learning rate of 0.001. We have used a batch size 50, the number of epochs is up to 50, and the categorical cross-entropy as a loss function. Table II shows the results for our proposed model, and it gives good achievement in terms of accuracy metrics that achieve 61.2 and 21.3 in BLEU metrics. Since there are few models that use PathVQA dataset for training, our model increases the accuracy by 2.2% and BLEU by 2.1%. One of the reasons for such a significant drop in performance is the presence of a new technique of vision transformers.

Table II. Results Comparison

Methods	Accuracy	BLEU
CNN + LSTM + stacked attention network [27]	59.4	19.2
Proposed (DeiT+BERT)	61.2	21.3

## Acknowledgment

The authors extend their appreciation to the Deanship of Scientific Research at King Saud University for funding this work through research group No. RG-1441-502.

## IV. Conclusion

Health care has rapidly grown with different tools and techniques to improve patients care. One of these techniques is a medical image. In this work, we propose full transformers for answering corresponding questions based on medical images. In particular, we use encoder modality for each input. The DeiT model is used to extract image features with a CNN teacher distilled model. The text feature extraction uses the BERT model by adding token, segment, and position embeddings layers. The answer was to predict the sequence by a decoder. Our model gets results of 61.2 accuracy score and 21.3 BLEU score. Empirical evaluation on the recently published benchmark dataset PathVQA shows that our approach achieves superior performance compared with the state-of-the-art Med-VQA model. In future work, we plan to explore a better evaluation strategy for evaluating the model. We also plan to introduce better individual models to handle each of the leaf node tasks.

## References

[1] E. Bercovich and M. Javitt, "Medical Imaging: From Roentgen to the Digital Revolution, and Beyond," *Rambam Maimonides Med. J.*, vol. 9, p. e0034, Oct. 2018, doi: 10.5041/RMMJ.10355.

[2] J. Ker, L. Wang, J. Rao, and T. Lim, "Deep Learning Applications in Medical Image Analysis," *IEEE Access*, vol. 6, pp. 9375–9389, 2018, doi: 10.1109/ACCESS.2017.2788044.

[3] M. M. A. Rahhal, Y. Bazi, H. AlHichri, N. Alajlan, F. Melgani, and R. R. Yager, "Deep learning approach for

active classification of electrocardiogram signals," *Inf. Sci.*, vol. 345, pp. 340–354, Jun. 2016, doi: 10.1016/j.ins.2016.01.082.

[4] C. Eickhoff, I. Schwall, and H. Muller, "Overview of ImageCLEFcaption 2017 – Image Caption Prediction and Concept Detection for Biomedical Images," p. 10.

[5] A. Qayyum, S. M. Anwar, M. Awais, and M. Majid, "Medical image retrieval using deep convolutional neural network," *Neurocomputing*, vol. 266, pp. 8–20, Nov. 2017, doi: 10.1016/j.neucom.2017.05.025.

[6] S. A. Hasan, Y. Ling, O. Farri, J. Liu, H. Muller, and M. Lungren, "Overview of ImageCLEF 2018 Medical Domain Visual Question Answering Task," p. 8.

[7] M. Sarrouti, A. Ben Abacha, and D. Demner-Fushman, "Visual Question Generation from Radiology Images," in *Proceedings of the First Workshop on Advances in Language and Vision Research*, Online, 2020, pp. 12–18. doi: 10.18653/v1/2020.alvr-1.3.

[8] X. Yan, L. Li, C. Xie, J. Xiao, and L. Gu, "Zhejiang University at ImageCLEF 2019 Visual Question Answering in the Medical Domain," p. 9.

[9] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," *ArXiv181004805 Cs*, p. 13, May 2019.

[10] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," *ArXiv14091556 Cs*, Apr. 2015, Accessed: Feb. 09, 2021. [Online]. Available: <http://arxiv.org/abs/1409.1556>

[11] M. Lin, Q. Chen, and S. Yan, "Network In Network," *ArXiv13124400 Cs*, Mar. 2014, Accessed: Jun. 12, 2021. [Online]. Available: <http://arxiv.org/abs/1312.4400>

[12] G. Chen, H. Gong, and G. Li, "HCP-MIC at VQA-Med 2020: Effective Visual Representation for Medical Visual Question Answering," p. 8.

[13] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2016, pp. 770–778. doi: 10.1109/CVPR.2016.90.

[14] B. Zhou, Q. Cui, X.-S. Wei, and Z.-M. Chen, "BBN: Bilateral-Branch Network With Cumulative Learning for Long-Tailed Visual Recognition," in *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, WA, USA, Jun. 2020, pp. 9716–9725. doi: 10.1109/CVPR42600.2020.00974.

[15] F. Ren and Y. Zhou, "CGMVQA: A New Classification and Generative Model for Medical Visual Question Answering," *IEEE Access*, vol. 8, pp. 50626–50636, 2020, doi: 10.1109/ACCESS.2020.2980024.

[16] Y. Khare, V. Bagal, M. Mathew, A. Devi, U. D. Priyakumar, and C. V. Jawahar, "MMBERT: Multimodal BERT Pretraining for Improved Medical VQA," *ArXiv210401394 Cs*, Apr. 2021, Accessed: May 31, 2021. [Online]. Available: <http://arxiv.org/abs/2104.01394>

[17] M. H. Vu, T. Lofstedt, T. Nyholm, and R. Sznitman, "A Question-Centric Model for Visual Question Answering in Medical Imaging," *IEEE Trans. Med. Imaging*, vol. 39, no. 9, pp. 2856–2868, Sep. 2020, doi: 10.1109/TMI.2020.2978284.

[18] A. Vaswani *et al.*, "Attention Is All You Need," *ArXiv170603762 Cs*, Dec. 2017, Accessed: May 31, 2021. [Online]. Available: <http://arxiv.org/abs/1706.03762>

- [19] A. Dosovitskiy *et al.*, “An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale,” *ArXiv201011929 Cs*, Oct. 2020, Accessed: May 31, 2021. [Online]. Available: <http://arxiv.org/abs/2010.11929>
- [20] Y. Bazi, L. Bashmal, M. M. A. Rahhal, R. A. Dayil, and N. A. Ajlan, “Vision Transformers for Remote Sensing Image Classification,” *Remote Sens.*, vol. 13, no. 3, p. 516, Feb. 2021, doi: 10.3390/rs13030516.
- [21] H. Touvron, M. Cord, M. Douze, F. Massa, A. Sablayrolles, and H. Jégou, “Training data-efficient image transformers & distillation through attention,” *ArXiv201212877 Cs*, Jan. 2021, Accessed: May 31, 2021. [Online]. Available: <http://arxiv.org/abs/2012.12877>
- [22] G. Hinton, O. Vinyals, and J. Dean, “Distilling the Knowledge in a Neural Network,” *ArXiv150302531 Cs Stat*, Mar. 2015, Accessed: Jun. 12, 2021. [Online]. Available: <http://arxiv.org/abs/1503.02531>
- [23] L. Bashmal, Y. Bazi, M. M. Al Rahhal, H. Alhichri, and N. Al Ajlan, “UAV Image Multi-Labeling with Data-Efficient Transformers,” *Appl. Sci.*, vol. 11, no. 9, p. 3974, Apr. 2021, doi: 10.3390/app11093974.
- [24] H. Tan and M. Bansal, “LXMERT: Learning Cross-Modality Encoder Representations from Transformers,” in *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, Hong Kong, China, Nov. 2019, pp. 5100–5111. doi: 10.18653/v1/D19-1514.
- [25] J. Lu, D. Batra, D. Parikh, and S. Lee, “ViLBERT: Pretraining Task-Agnostic Visiolinguistic Representations for Vision-and-Language Tasks,” *ArXiv190802265 Cs*, Aug. 2019, Accessed: May 31, 2021. [Online]. Available: <http://arxiv.org/abs/1908.02265>
- [26] R. Hu and A. Singh, “UniT: Multimodal Multitask Learning with a Unified Transformer,” *ArXiv210210772 Cs*, Mar. 2021, Accessed: May 31, 2021. [Online]. Available: <http://arxiv.org/abs/2102.10772>
- [27] X. He, Y. Zhang, L. Mou, E. Xing, and P. Xie, “PathVQA: 30000+ Questions for Medical Visual Question Answering,” *ArXiv200310286 Cs*, Mar. 2020, Accessed: Jun. 14, 2021. [Online]. Available: <http://arxiv.org/abs/2003.10286>
- [28] M. Malinowski and M. Fritz, “A Multi-World Approach to Question Answering about Real-World Scenes based on Uncertain Input,” *ArXiv14100210 Cs*, May 2015, Accessed: Jun. 12, 2021. [Online]. Available: <http://arxiv.org/abs/1410.0210>
- [29] K. Papineni, S. Roukos, T. Ward, and W.-J. Zhu, “Bleu: a Method for Automatic Evaluation of Machine Translation,” in *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics*, Philadelphia, Pennsylvania, USA, Jul. 2002, pp. 311–318. doi: 10.3115/1073083.1073135.

# Deep Learning for Face Detection and Recognition

Tuba Elmas ALKHAN  
Department of Computer Engineering  
Istanbul Sabahattin Zaim University  
Istanbul, Turkey  
karabala.masa@std.izu.edu.tr

Alaa Ali Hameed  
Department of Computer Engineering  
Istanbul Sabahattin Zaim University  
Istanbul, Turkey  
0000-0002-8514-9255

Akhtar Jamil  
Department of Computer Engineering  
Istanbul Sabahattin Zaim University  
Istanbul, Turkey  
0000-0002-2592-1039

**Abstract**— With the more development of machine learning and deep learning, face recognition technology based on convolutional neural network (CNN) has become the most necessary and used methodology within the field of face recognition. A face recognition model could be a technology capable of distinctive a personal from a picture or a video. Various strategies for face recognition systems are effective, and they work by comparing selected facial features from images with faces in the database. This paper creates a system that uses Convolutional Neural Networks (CNN) to acknowledge students' emotions from their faces. We achieved an accuracy of 74.41% and validation accuracy of 77.00% on the fer2013 dataset to classify seven different emotions through facial expressions.

**Keywords**—Face Detection, Face Recognition, Deep Learning, Emotion recognition, Convolutional neural networks (CNN), Student facial expression.

## I. Introduction

Face emotion recognition is an energetic and vital area of research. Especially these days, due to the spread of the COVID-19 epidemic, it become distance education. These systems play an essential role in our daily life and make it much more manageable. Face emotion recognition has been implemented in medicine, psychology, interactive games, public security and distance education, etc.

Face recognition in videos is challenging due to variations in pose, illumination, or facial expression. But this is an important task that has been widely utilized in several practical applications like security monitoring, surveillance [1], etc.

The face is that the most important part of the body. It is vital and expressive. It can transfer many emotions silently. Facial expression recognition determines an emotion from face images. Generally, six basic emotions are categorized, which are the same across all cultures; happiness, sadness, surprise, angry, fear, disgust, neutral [1].

CNN has been proven to be very active for various computer vision works, such as object or face detection and classification [2]. Applying facial expression recognition system to the field of education can allow to detect, capture and record the emotional changes of students during the learning process and supply higher reference for academics to teach based on students' abilities. [3]

The facial recognition system involves two steps: face detection, which identifies human faces in images. In contrast, face recognition matches the face from a video or an image against a database of faces to recognize it. Both are similar but

different have different objectives. Researchers have proposed several facial detection and recognition systems, which will be discussed in detail in the next section.

The purpose of this article is to execute emotion recognition in education area through a system using convolutional neural network (CNN) that analyzes facial expressions of students. CNN is a deep learning algorithm used for image classification. It includes of several stages image processing for extracting feature representation. There are several deep learning methods for extracting more complex features, such as Autoencoders, Recurrent Neural Network, Gradient Descent and Convolutional Neural Networks.

This study implements an automated system to realize emotion recognition in education field. The system analyzes student facial expressions and gives feedback to an educator with using Convolutional Neural Network. Several classification algorithms were applied to learn instant emotional state (Random Forest, Artificial Neural Network (ANN), Support Vector Machine (SVM), K-Nearest Neighbor (KNN) and Classification & Regression Trees). E-learning has several advantages, such as saving time and money. Through this learning, All of students can uses the contents in anytime and anywhere, which leads to good participation, retention, scalable and offers personalization, but it does not provide enough face-to-face interactivity between an educator and learners.

## II. LITERATURE REVIEW

This section highlights some developments made in the field of facial emotion recognition in various areas such as medicine, health, psychology, online education and biomedical engineering . Today, Face emotion recognition is a vital and important area of research. Especially these days, due to the spread of the COVID-19 epidemic, it was distance education. The detection of facial emotions is possible in online education. Therefore, it can facilitate academics to adjust their performance depending on the students' emotions. In general, Artificial intelligence includes deep learning and machine learning. Many machines learning and deep learning algorithms are used in this field. Convolutional neural Networks (CNNs) have become the most important and used method in the field of face recognition. Which is a deep learning algorithm used for image classification. It includes of several stages image processing for extracting features representation.

In [6], the authors propose a CNN architecture called Trunk Branch Ensemble Convolutional Neural network

(TBE-CNN) to overcome problems in facial recognition from a video. They used this system in surveillance. Surveillance applications need to be capable of detecting and recognizing faces quickly. This system must be able to withstand changes in blur, zoom, illumination and different poses. This version extracts features effectively via way of means of sharing the low- and middle-degree convolutional layers.

In [7], a new system for face recognition based on Stacked Automatic Convolutional Encoder (SCAE) and sparse representation was presented. The system can extract deeper and abstract features with high recognition speed. However, the recognition rate is not high, so they need to develop the system.

Viola-Jones framework has been widely used by researchers Padilla and Costa for detecting the location of faces, this work focuses on the appraisal of face detection classifiers, such as OpenCV. The system needs images with and without faces (positive & negative pictures) for training the classifier and extract features (Haar) from images. The authors evaluated the performance of some classifiers and tested their accuracy [8].

Authors in [2] proposed a school system using convolutional neural network for helping professors and instructors to change their academic performance based on students' emotions. First, they detect students faces from an image by using Haar Cascades Classifiers, then emotion recognition by using CNN with seven types of expressions. The system achieved an accuracy of 70% on fer-2013 dataset.

In [9], authors developed a system that determines students emotions and provides feedback to improve distance education and update learning content. Head and eyes movement can help to comprehend student attention and concentration level. The system is suitable and active for detecting students' negative emotions. Authors discussed some face detection algorithms such as Local Binary pattern (LBP), Neural network and Ada Boost.

(Chang et al., 2018) designed a new Convolutional Neural Network based on ResNet to extract features and Complexity Perception Classification (CPC) algorithm for facial expression recognition using three different classifiers (Softmax, SVM, Random Forest). It improved the recognition accuracy and fixed some misclassified expression categories. CNN and Softmax with CPC algorithm has achieved accuracy 71.35% for Fer2013 and 98.78% for CK+ [16].

(Jiang et al., 2020) Gabor Convolutional Network is shown in three types of datasets (real world affective faces, FER+ and Fer2013). The proposed approach includes four Gabor convolutional layers with two fully connected layers. They find the optimal model by changing the numbers of layers and the numbers of units at the convolutional layers. Then they discussed and compared the proposed GCN model with different models such as AlexNet and ResNet. The GCN has achieved best accuracy on the fer2013 dataset [15].

In their study, (Ayvaz et al.) discussed a new facial emotion recognition system with the help of several algorithms such as (Random Forest, Support Vector Machine (SVM), K-Nearest Neighbor (KNN) and Classification & Regression Trees) the system can classify the emotions of the students. The system detects facial emotional of the students and gives response to an instructor according to the facial

expression of the learner. SVM provides the best prediction accuracy rate of 98.24% [10].

Recently, many works [12, 2] used CNN for facial expressions recognition. The recognition of human facial expressions is a hard problem for deep learning and machine learning, so the convolutional neural network is used to conquer the problems in facial expression classification.

In their study, Roman RADIL et al. The performance of the proposed Convolutional Neural Network with three image recognition methods like Local Binary Patterns Histograms (LBPH), K Nearest Neighbor (KNN) and Principal Component Analysis (PCA) is tested. The result shows that the Local Binary Patterns Histograms provide better results than Principal Component Analysis and K-Nearest Neighbor, and an accuracy rate of 98.3% was achieved for proposed CNN [11].

Saravanan et al. discussed the Classification of images of human faces into one of 7 basic emotions (Fear, Disgust, Surprise, Anger, Sadness and Happiness), authors proposed approach (CNN) Convolutional Neural Network model which content of six convolutional layers, two max pooling layers and tow fully connected layers. This model achieved a final accuracy of 0.60.

In [14], the authors created a model using CNN to detect facial expressions in real time using a webcam. The model is used to classify the expression of human faces, and the model gave a training accuracy of 79.89% and a test accuracy of 60.12%.

(Wang et al., 2020) Online education has developed Because of the spread of the COVID-19 pandemic, which has led to the closure of schools and the transfer of education to distance education .so author proposed a system combining a Face Emotion Recognition (FER) algorithm and online courses platforms depend on the architecture of CNN [17].

### III. MATRIALS AND METHODOLOGY

#### A. Dataset

For our deep learning model to be good and smart to discover expressions, we need to train it with a facial expression dataset. Here we used FER-2013 dataset. FER-2013 dataset is an open-source dataset to recognize facial expression, which was shared on Kaggle through the ICML 2013 conference. The dataset contains of 35.887 grayscale images with 48x48 sized of face images, divided into 3.589 test and 28.709 train images. The dataset contains of facial expressions belonging to these seven emotions (Happy, Sad, Neutral, Surprise, Fear, Angry and Disgust). Figure 1 shows some example images from the FER-2023 dataset, Table 1 illustrates the description of the dataset. The image dataset consists of grayscale images, and we kept size the same for our training and testing (300 x300).

Table I. FER2013 dataset description

Label	Number of images	Emotion
0	4593	Angry
1	547	Disgust
2	5121	Fear

3	8989	Happy
4	6077	Sad
5	4002	Surprise
6	6198	Neutral



Figure I. Sample images from the FER2013 dataset

### B. Proposed Method

In this part, we will describe our proposed system that uses Convolutional Neural Network (CNN) model to analyze the students' facial expressions. The first step of our system is to detect the face from images or video (video is a set of images), then use these face images as input to the network. Lastly, by using CNN the system classifies the expression of a student's face into one of these (happy, sad, fear, angry, surprise, neutral, disgust) expressions.

Convolutional neural network is a type of artificial neural network that uses a convolution method for extracting the large number of features from the input data.

CNN model contains 3 types of main layers: Convolutional layer, pooling layer and fully connected layer. Figure II shows the CNN architecture.

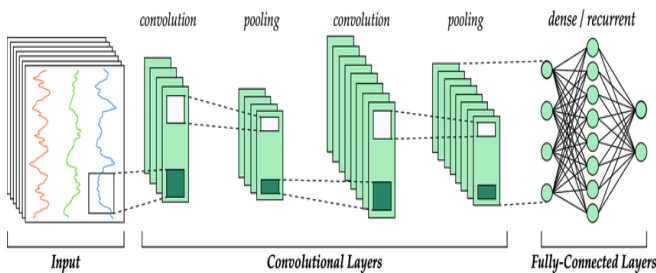


Figure II. Convolutional Neural Network Architecture

**Convolutional Layer:** We used convolutional layer to extract the different attributes from an entire image. The convolution keeps a spatial association among pixels by learning features, then the images will be convoluted by use a group of learnable neurons. This creates in the output picture a feature map that provides us some information about this image. Finally, the feature maps are fed to the next layer for learning more features. We can explain the convolution in other words, multiplied two images that can be represented as a matrix to get an output that is used to extract a large number of features from the image.

The convolution formula is represented in Equation 1:  $f$  is the input photo,  $*$  is for the convolution operation and  $g$  was the filters matrix. Figure III shows the convolution operation.

$$f(x, y) * g(x, y) = \sum_{n_1=-\infty}^{\infty} \sum_{n_2=-\infty}^{\infty} f(n_1, n_2) \cdot g(x - n_1, y - n_2) \quad (1)$$

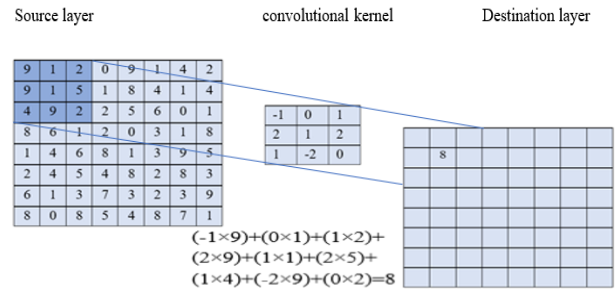


Figure III. Convolution operation.

**Pooling Layer:** Pooling is also known as down sampling or subsampling. Pooling layer can down sampling of any features map yet holds the significant data. There are three common pooling methods: Sum pooling, Max pooling and Average pooling. The most typically approach used in pooling is that Max pooling. Max pooling is used to gradually minimize the spatial size of the input, it controls overfitting (regularization) and Invariance to small translations of the input). Pooling layer provides higher generalization, resistant to distortion and quicker convergence. It is typically positioned among the convolutional layers. Figure IV shows an example of Max Pooling operation.

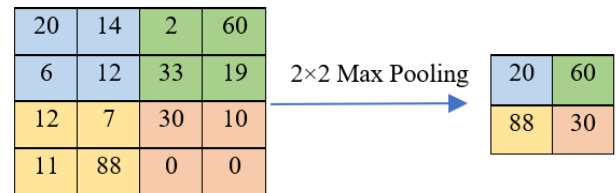


Figure IV. Max pooling layers.

**Fully connected layer:** Fully connected layer in a neural network is that layer in which all inputs of one layer are connected to each neuron of the next layer. The intention of using the FCL is to apply the output of the previous layers like convolutional and pooling layers to classify the input image into different classes according to the training dataset. The term fully connected layer means that all filters of the previous layer are linked to all filters of the next layer. Fully connected layers are positioned earlier than the convolutional neural network sorting output and used to flatten results before classification. In short, the convolutional layers and pooling layers work as characteristic extractors from the inputs and the fully connected layers work as classifier.

### IV. EXPERIMENTAL RESULTS

We designed our CNN model. Here we used FER-2013 dataset. It is an open-source dataset, shared on Kaggle. The dataset includes seven categories such as Happy, Sad, Fear, Angry, Disgust, Neutral and Surprise. The training set consists

of about 17084 images. The testing set consists of about 4180 images. We will be dealing with 3 classes (Happy, Sad, Neutral) in training our model. image rows= 48, image column= 48, determine the size of the image matrix which we will be feeding to our model. Batch size= 32, batch size is the number of samples that are processed before updated the model. The number of complete passes through the training dataset was 25 epochs. We used ImageDataGenerator class to expand the size of a training dataset. ImageDataGenerator class allowed us to expand the training pictures by using some arguments like rescale, rotation, zoom, horizontal or vertical flip and shifts. We generate a sequential model. We designed our Convolutional Neural Networks (CNN) model contains of 8 convolutional layers, 4 max pooling layers and 3 fully connected layers at the end Softmax classifier with seven categories namely happy, sad, surprise, neutral, fear, disgust and angry is used to classify the given image input. There are 7 types of layers which we have used. In the first convolutional layer the number of filters were 32, filter size of (3, 3), padding = 'same' and with a kernel initializer = he\_normal. We are using this layer to extract various attributes from the enter picture, and a convolution saves a spatial connection among pixels by learning features using small squares of input data. In our CNN architecture, the Rectified linear unit (ReLU) activation function has been used. ReLU is that the most generally used activation function. It is mainly applied in hidden layers of the Neural network.

$$R(z) = \max(0, z) \quad (2)$$

The equation for ReLU activation function is as shown above. If Z is positive z, otherwise output will be 0. ReLU is a smaller amount computationally costly than tanh and sigmoid, as it entails less complicated mathematical operations. Solely some neurons are activated at a time, which makes the network active and easy to calculate. Batch normalization layer batch normalization accelerates training, provides some regularization and reducing generalization error. Then we have added max pooling layer to decrease the dimensionality of each feature map. Here we have used the pooling size as (2,2). We used a dropout layer to reduce the overfitting. We used Dropout as 0.5, which means half of the neurons will be ignored. Flatten layer used to convert the pooled feature map to one column or a vector. We used this layer to flatten the input or the output of the preceding layers. Dense layer (Fully Connected Layer) the aim of this layer is to utilize the output of the convolutional and pooling layers to classify the input images into different classes according to the training dataset. Here we used 64 neurons and the kernel initializer he\_normal. At the end, the Softmax Classifier is used to classify human faces. In the output layer we used Softmax as an activation function to produce probabilistic output for each class. Softmax takes a vector consisting of N real numbers, and then normalizes that vector into a range of values between (0, 1). Softmax conversion the input value which may be negative, positive, zero or larger than one, values range from zero to one, so they can be interpreted as probabilities. Figure V shows the structure of the proposed model.

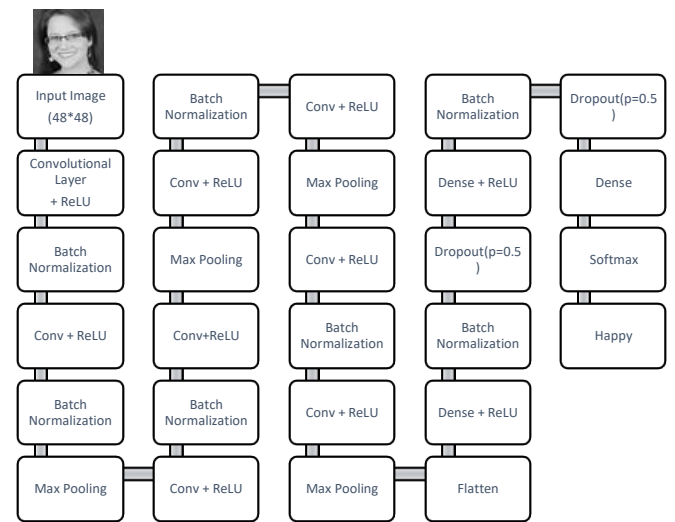


Figure V. The structure of the proposed model.

There are many approaches to recognize faces in an image, however we are going to use the Haar Cascade Classifiers that are machine learning models trained to find and detect some features such as faces, eyes and lips in an image. If the picture includes these features that means the picture contains a face, otherwise there is no face in the photo.

We tested our model on 3 optimizers such as SGD, Adadelta and Adam. The convolutional neural network consisting of 8 convolutional layers was trained on 17084 images and validated on 4180 images using adaptive moment estimation (Adam) optimizer, at 25th epoch, the batch size of 32 and learning rate of 0.001 the training accuracy was 74.41%. Validation accuracy of 77.00%. Whereas training loss was 0.5945 and validation loss was 0.5493, and we calculated confusion matrix. It can be seen from the graph Figures VI, VII and VIII.

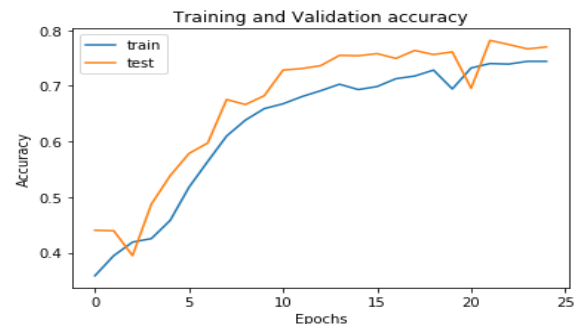


Figure VI. Accuracy over train and test data in our CNN with Adam optimizer.

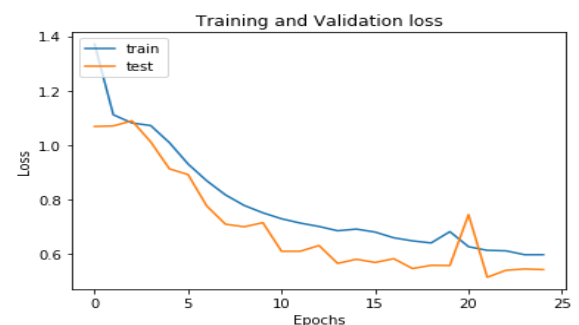


Figure VII. Loss over test and train data with Adam optimizer.



Figure VIII. Confusion matrix of proposed method with Adam optimizer.

Using the Stochastic Gradient Descent optimizer, the learning rate of 0.001, the batch size of 32 and 18 epochs lead to a low accuracy of 0.5267 and validation accuracy of 0.5507. However, upon setting the learning rate to 0.1 and 40 epochs the highest accuracy of 0.7236 and validation accuracy of 0.7683 was attained.

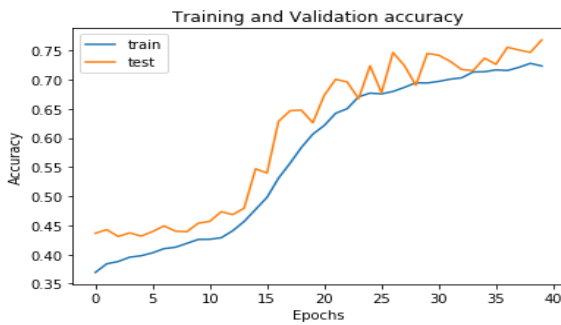


Figure IX. Accuracy over test and train data with SGD optimizer in the proposed CNN.

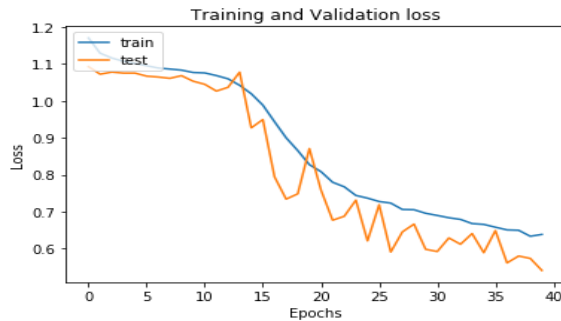


Figure X. Loss over test and train data with SGD optimizer.

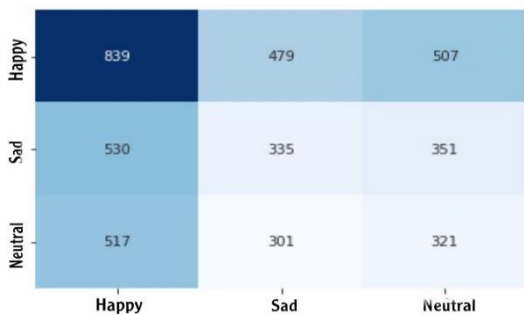


Figure XI. Confusion Matrix of proposed method with SGD optimizer.

The Adadelta optimizer gave an accuracy of 0.7179 and validation accuracy of 0.7563 was attained over 50 epochs, the learning rate was 0.1 and the batch size was 32.

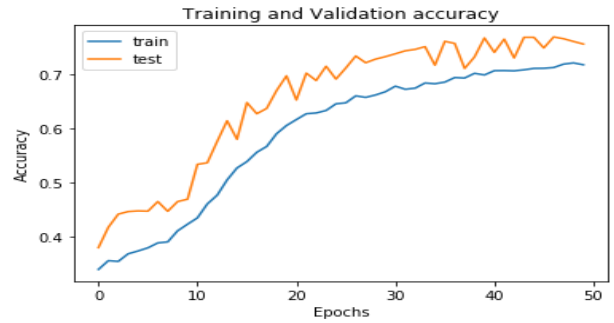


Figure XII. Accuracy over test and train data with Adadelta optimizer in proposed CNN.

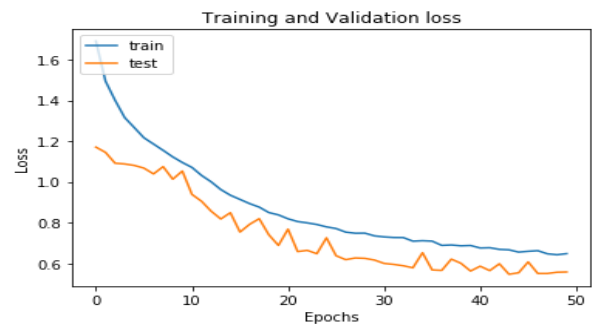


Figure XIII. Loss over train and test data with Adadelta optimizer in CNN.



Figure XIV. Confusion Matrix with Adadelta optimizer using the proposed method.

Based on these results, it is able to be concluded that the Adam optimizer give us the best accuracy. We used these functions EarlyStopping() and ModelCheckpoint() in Keras to get better result in our model. In machine learning, particularly in deep learning, early stopping is one of the most widely used regularization techniques to avoid the overfitting problem in neural network. Early stopping is a technique that permits you to halt training once the model performance stops enhancing at the validation dataset. This function can stop the execution early. This is done by checking some parameters. Monitor: allows to specify the performance measure to monitor. Here, we are tracking the validation loss. min\_delta: is a threshold to whether or not quantify a loss at an epoch as improvement or not. If the change of loss is under than the min\_delta, will depend on as no improvement. We've got given it zero.

Patience: represents the quantity of epochs which training will be stopped after it because no improvement and the loss start to increase. We have given patience 10. Verbose: To discover and print the training epoch on which training was stopped, verbose can be set to 1. restore\_best\_weights: whether to retrieve weights with the best value of the monitored quantity. Here i have given it True.

The ModelCheckpoint class allows you to define where to checkpoint the model weight to keep it. Therefore, the weights may be loaded later to carry on training from the saved state. We monitored the validation loss and minimized the loss using the mode='min' parameter.

Our system realizes the faces of the input images of the students by using Haar cascades detector then classifies them into one of seven basic expressions. The proposed method achieved an accuracy of 77% using Adam optimizer on FER2013 dataset at the 25 epochs.

## V. CONCLUSION

In this study, our aim was to detect the face then to classify facial expressions, so we present a convolutional neural network model for the recognition of the facial expressions of the students. With the help of deep learning and machine learning technologies we can Classification of the emotions of the online learner, therefore it is able to help the instructor to recognize students' understanding during his presentation and gives feedback to an educator. In our future work we are going to focus on applying Convolutional Neural Network model on 3D students' face image so as to extract their emotions.

## REFERENCES

- [1] "Face Recognition in Real-world Surveillance Videos with Deep Learning Method," pp. 239–243, 2017.
- [2] I. Lasri, "Facial Emotion Recognition of Students using Convolutional Neural Network," pp. 0–5, 2019.
- [3] R. Ranjan et al., "A fast and accurate system for face detection, identification, and verification," arXiv, vol. 1, no. 2. IEEE, pp. 82–96, 2018.
- [4] H. Zhang, A. Jolfaei, and M. Alazab, "A Face Emotion Recognition Method Using Convolutional Neural Network and Image Edge Computing," IEEE Access, vol. 7, pp. 159081–159089, 2019.
- [5] S. S. Mohamed, W. A. Mohamed, A. T. Khalil, and A. S. Mohra, "Deep Learning Face Detection and Recognition," no. June, 2019.
- [6] C. Ding and D. Tao, "Trunk-Branch Ensemble Convolutional Neural Networks for Video-Based Face Recognition," vol. 40, no. 4, pp. 1002–1014, 2018.
- [7] L. Chang, J. Yang, S. Li, H. Xu, K. Liu, and C. Huang, "Face Recognition Based on Stacked Convolutional Autoencoder and Sparse Representation," *Int. Conf. Digit. Signal Process. DSP*, vol. 2018–November, pp. 1–4, 2019.
- [8] R. Padilla, C. C. Filho, and M. Costa, "Evaluation of haar cascade classifiers designed for face detection," *J. WASET*, vol. 6, no. 4, pp. 323–326, 2012.
- [9] L. B. Krithika and G. G. Lakshmi Priya, "Student Emotion Recognition System (SERS) for e-learning Improvement Based on Learner Concentration Metric," *Procedia Comput. Sci.*, vol. 85, no. Cms, pp. 767–776, 2016.
- [10] U. Ayvaz, H. Gürüler, and M. O. Devrim, "Use of Facial Emotion Recognition in E-Learning Systems," *Inf. Technol. Learn. Tools*, vol. 60, no. 4, p. 95, 2017.
- [11] P. Kamencay, M. Benco, T. Mizdos, and R. Radil, "A new method for face recognition using convolutional neural network," *Adv. Electr. Electron. Eng.*, vol. 15, no. 4 Special Issue, pp. 663–672, 2017.
- [12] A. Fathallah, L. Abdi, and A. Douik, "Facial expression recognition via deep learning," *Proc. IEEE/ACS Int. Conf. Comput. Syst. Appl. AICCSA*, vol. 2017–October, no. October, pp. 745–750, 2018.
- [13] M. Mohammadpour, H. Khaliliardali, S. M. R. Hashemi, and M. M. Alyannezhadi, "Facial emotion recognition using deep convolutional networks," *2017 IEEE 4th Int. Conf. Knowledge-Based Eng. Innov. KBEI 2017*, vol. 2018–Janua, pp. 0017–0021, 2018.
- [14] I. Talegaonkar, K. Joshi, S. Valunj, R. Kohok, and A. Kulkarni, "Available on: Elsevier-SSRN Real Time Facial Expression Recognition using Deep Learning," 2019.
- [15] P. Jiang, B. Wan, Q. Wang, and J. Wu, "Fast and Efficient Facial Expression Recognition Using a Gabor Convolutional Network," *IEEE Signal Process. Lett.*, vol. 27, pp. 1954–1958, 2020.
- [16] T. Chang, G. Wen, Y. Hu, and J. J. Ma, "Facial expression recognition based on complexity perception classification algorithm," arXiv, 2018.
- [17] W. Wang, K. Xu, H. Niu, and X. Miao, "Emotion Recognition of Students Based on Facial Expressions in Online Education Based on the Perspective of Computer Simulation," *Complexity*, vol. 2020, 2020.



# A Lightweight and Interpretable Deepfakes Detection Framework

Muhammad Umar Farooq

Department of Software Engineering  
University of Engineering and Technology  
Taxila, Pakistan  
softwareengineerumar@gmail.com

Ali Javed

Department of Computer Science  
University of Engineering and Technology  
Taxila, Pakistan  
ali.javed@uettaxila.edu.pk

Khalid Mahmood Malik

Department of CS and Engineering  
Oakland University  
Rochester, MI, USA  
mahmood@oakland.edu

Muhammad Anas Raza

Department of Mechanical Engineering  
University of Engineering and Technology  
Taxila, Pakistan  
memanasraza@gmail.com

**Abstract**—The recent realistic creation and dissemination of so-called deepfakes poses a serious threat to social life, civil rest, and law. Celebrity defaming, election manipulation, and deepfakes as evidence in court of law are few potential consequences of deepfakes. The availability of open source trained models based on modern frameworks such as PyTorch or TensorFlow, video manipulations Apps such as FaceApp and REFACE, and economical computing infrastructure has eased the creation of deepfakes. Most of the existing detectors focus on detecting either face-swap, lip-sync, or puppet master deepfakes, but a unified framework to detect all three types of deepfakes is hardly explored. This paper presents a unified framework that exploits the power of proposed feature fusion of hybrid facial landmarks and our novel heart rate features for detection of all types of deepfakes. We propose novel heart rate features and fused them with the facial landmark features to better extract the facial artifacts of fake videos and natural variations available in the original videos. We used these features to train a light-weight XGBoost to classify between the deepfake and bonafide videos. We evaluated the performance of our framework on the world leaders dataset (WLDR) that contains all types of deepfakes. Experimental results illustrate that the proposed framework offers superior detection performance over the comparative deepfakes detection methods. Performance comparison of our framework against the LSTM-FCN, a candidate of deep learning model, shows that proposed model achieves similar results, however, it is more interpretable.

**Index Terms**—Deepfakes, Multimedia Forensics, Random Forest Ensembles, Tree boosting, XGBoost, Faceswap, Lip sync, Puppet Master.

## I. INTRODUCTION

Recent advancements in deep learning (DL) have impacted the way we solve complex technical problems in computer vision (CV) and robotics. With the widespread availability of video synthesis repositories and video manipulations Apps such as FaceApp [1] and REFACE [2], video manipulation has become easy, even for a layman. Video synthesis is beneficial in some ways like avatar creation, animated video content creation, etc. Sometimes videos are synthesized just for the sake of fun, like a recent realistic Tiktok video of Tom Cruise

[3]. However, the case is not always that simple. Depending on the time and context, deepfakes pose a serious threat to society. With deepfakes, celebrities are defamed, and election campaigns could be manipulated. DL based video synthesis tools use generative adversarial networks (GAN) under the hood. The adaptive nature of GAN made it difficult to develop a robust detection solution. Whenever a deepfakes detection model is developed, we witness some variant of a GAN based generation model to exploit the newly developed detection model by manipulating its cues. Thus, deepfakes creation and detection is a constant battle between the ethical and unethical machine learning (ML) experts.

Deepfakes detection got much attention in the last decade after realistic fake videos of politicians and celebrities got viral via social media platforms. Current deepfake videos are categorized as face-swap, lip-sync, and puppet master [4]. In face-swap deepfakes, face of a target person is added at the place of a source person in the original video to create a fake video of the target person. In lip-sync deepfakes, lips of a person are synced for an audio to reflect that person is speaking the text in that audio. In puppet-master, the face of the target person is placed in the original video but facial expressions of the source person are retained on the target face to make the fake more realistic. Most of the existing detection solutions target specific types of deepfakes, however, generic solutions capable of countering all types of deepfakes are less explored. For example, Agarwal et al. [5] proposed a detection technique for lip-sync deepfakes. This technique exploited the inconsistencies between the viseme (mouth shape) and a phoneme (spoken word). This work applied manual and CNN based techniques to compute the mapping of viseme to phoneme. This model is good for a specific set of seen data. However, model performance can degrade on unseen data for different patterns of viseme to phoneme mapping, with the change of speaking accent or even non-alignment of audio-to-video.

Most of the existing systems are unable to perform well on

all three types of deepfakes. Moreover, deepfakes detection models based on the traditional classifiers like SVM, works only where data is linearly separable. CNN based models are computationally more complex and are black-box in terms of prediction. Therefore, this paper addresses the following research questions:

- 1) Is it possible to improve the detection accuracy of deepfakes using hybrid landmark and heart-rate features on a diverse dataset containing all three types of deepfakes?
- 2) Is it possible to create a generalized detection model based on proposed hybrid landmark and heart-rate features and ensemble learning?
- 3) Is it possible to achieve the same accuracy as deep learning models but improve the interpretability by using an ensemble of supervised learning?

Existing deepfake detection techniques are broadly categorized as handcrafted features [6]–[9] based or DL based [10]–[14]. For example, Yang et al. [9] used 68-D facial landmark features to train an SVM classifier for detection. This work achieved good performance on good quality videos of UADFV [9] and DARPA MediFor [15] datasets but was unable to perform well on low quality videos. Moreover, the evaluation of this work did not consider all types of deepfakes. Matern et al. [6] used 16-D texture based eyes and teeth features for the exploitation of the visual artifacts to detect video forgeries like face-swap and Face2Face. Most important aspect of this work was to detect the difference in eye color of a POI for detection of face-swap deepfakes detection by exploiting the missing details like reflection in eye color. Additionally, this work uses face border and nose tip features along with eye color features for Face2Face deepfakes detection. This technique [6] has a limitation of working only for faces with clear teeth and open eyes. Lastly, the evaluation of this work was only performed on FaceForensics++ [10] dataset. Li et al. [7] used the targeted affine warping artifacts introduced during deepfakes generation. Targeting the specific artifacts reduced the training overhead and improved the efficiency. However, these specific artifacts selection can compromise the robustness of this technique by making it difficult to detect a deepfake with slightly new transformation artifacts. Agarwal et al. [8] used an open source toolkit OpenFace2 [16] for facial landmark features extraction. Some features were derived based on extracted landmark features. These derived features were then used along with action unit (AU) features to train a binary SVM for deepfakes detection. This technique was proposed for five POIs where all POIs were linearly separable in a t-SNE plot. However, for an increased number of POIs in the updated dataset [17], performance of this technique was significantly degraded. In their extended work, Agarwal et al. [18] proposed a framework based on spatial and temporal artifacts in deepfakes. This framework is based on some threshold based rules to classify a video as real or fake. This rule-based approach would work on selected datasets, however, performance of this hard coded threshold oriented approach is expected to degrade on unseen data. In [18],

authors proposed a new framework, ‘FakeCatcher’, which uses biological signals from three face regions in the real videos to detect the fake videos. FakeCatcher applied many transformations on biological features like autocorrelation, power spectral density, wavelet transform, etc. Authenticity decision is based on the aggregated probabilities of two probabilistic classifiers (SVM and CNN). Performance was evaluated on their own customized dataset, however, it is not evaluated on all three types of deepfakes.

Besides the handcrafted features-based methods, deep learning-based methods are also being employed for deepfakes detection. Guera et al. [10] applied a DL based technique to detect the deepfakes. This technique applied a CNN to extract features followed by a long-short term memory (LSTM) to learn those features. Important contribution of this work was the exploitation of temporal inconsistencies among deepfakes for classification. However, this approach is unable to identify all three types of deepfakes. Afchar et al. [11] designed a neural network (MesoNet) to detect deepfakes and Face2Face video forgeries. This work designed an end-to-end architecture with convolutional and pooling layers for feature extraction followed by dense layers for classification. These methods [10], [11] were evaluated on videos collected from random websites rather than a standard dataset that doubted the robustness of these approaches for a large-scale and diverse standard dataset. Nguyen et al. [12] designed a capsule network to expose multiple types of tampering in images and videos. This framework aimed at detection of face swapping, facial re-enhancements and computer generated images. This framework used dynamic routing and expectation-maximization algorithms for performance improvement. The Capsule network employed the VGG-19 for latent face features extraction and used them for classification of original and bonafide videos. Framework is good at detecting face-swap forgeries in FaceForensics dataset, however, not evaluated on lip-sync and puppet-master deepfakes and complex in terms of computations. Sabir et al. [13] proposed a method based on DL to feed cropped and aligned faces to a CNN (ResNet and DenseNet) for feature extraction followed by an RNN for classification. Most important aspect of this work was to use features from multiple levels of CNN to incorporate mesoscopic level features extraction. This work [13] only used FaceForensics++[11] dataset for evaluation and didn’t consider lip-sync and puppet-master deepfakes. Yu et al. [14] used a CNN to capture the fingerprints of GAN generated images to perform the classification of synthetic and real images. This technique targeted fake images generated with four GAN variants ProGAN, SNGAN, CramerGAN, MMDGAN, but might not be able to detect fake images generated with a new GAN variant. In [19], authors used an ensemble of four CNNs to achieve good results on DFDC. An attention mechanism was added to EfficientNetB4 to get the insights of the training process. EfficientNetB4 and EfficientNetB4Att were trained as end-to-end training, whereas, EfficientNetB4ST and EfficientNetB4AttST were trained in Siamese training settings. Important aspect of this method was the data augmentation

(i.e. down sampling, hue saturation, JPEG compression, etc.) during training and validation for model robustness. Moreover, this technique performs well on large face-swap dataset DFDC but not evaluated on all three types of deepfakes and is computationally complex. In [20], authors used EfficientNet (a CNN) and gated recurrent unit (GRU) (an RNN) to exploit spatiotemporal features in the video frames to detect deepfake videos. This work included data augmentation on real videos during training to balance classes as DFDC is highly class imbalanced. Moreover, this architecture performs well on large face-swap dataset DFDC but is not evaluated on all three types of deepfakes and is complex in terms of computations.

Most methods based on handcrafted features [6]–[9] fail to generalize well on different types of deepfakes like lip-sync and puppet-master. CNN based techniques [10]–[14] are computationally complex and black-box in terms of generating the output. Moreover, these methods exploit some GAN specific artifacts produced during generation. So, they might fail to detect deepfakes, generated with a new GAN architecture.

To address the above mentioned problems and limitations of existing works, this paper proposes a lightweight model based on feature fusion of facial landmarks and heart rate features. For landmark features, we analyzed the impact of each landmark features category before final features selection. We analyzed the impact of different combinations of features categories. We started with two most effective features categories and then added one category in the feature-set at a time in the decreasing order of effectiveness. We disregarded the concept of the POIs being linearly separable, because that concept becomes invalid with a higher number of POIs. We used the XGBoost [21] for classification purposes. XGBoost uses Bagging in Random Forest for variance related errors and gradient boosting algorithm for bias related errors. XGBoost successfully addresses the data classification problem where data points are not linearly separable.

The main contributions of this paper are as follows:

- We propose a lightweight and interpretable deepfakes detection framework capable of accurately detecting all types of deepfakes namely, faceswap, puppet-master and lipsync.
- We propose novel heart rate features and fused them with a robust set of selected facial landmark features for deepfakes detection.
- We highlight that an XGBoost based solution is lightweight as compared to CNN based solutions and better generalize as compared to other conventional classifiers like SVM, KNN, etc.

Rest of the paper is structured as follows. Section 2 presents the details of feature engineering and model development. In Section 3, we present the details of performance evaluation and comparative analysis w.r.t to state of the art methods. Finally, we conclude our paper in Section 4.

## II. METHODOLOGY

This section provides an overview of the proposed framework. As shown in the Figure I, the input video is processed

to extract 850-D facial landmarks and 63-D heart rate features. XGBoost classifier is used for classification. Classifier is trained on each sub-category of landmark and heart rate features. Finally, we reduce the dimensions of our features to select the most reliable features among all to make the final features-set. XGBoost classifier is trained on the final features-set to classify the video as fake or bonafide. The process flow of the proposed solution is shown in Figure I.

### A. Features Extraction

Effective features extraction is crucial for any classification task. For this purpose, we proposed a fused features-set consisting of our novel heart rate features and the facial landmark features. We extracted facial Landmark features using the OpenFace2 [21] toolkit. For heart rate features, we selected seven regions of interest as shown in Figure I. Seven ROIs are right cheek (RC), left cheek (LC), chin (C), forehead (F), outer right edge (OR), outer left edge (OL), and center (C). We calculated RGB values of all ROIs and then applied some transformations to create heart rate features. Details of transformations are as follows:

$$HR_s = \{Z_R, Z_G, Z_B\} \quad (1)$$

$$HR_r = \{Z_R/Z_G, Z_R/Z_G, Z_G/Z_B\} \quad (2)$$

Where  $HR_s \in \{RC, LC, C, F, OR, OL, CE\}$   
 $\& R \leftrightarrow red, G \leftrightarrow green, B \leftrightarrow blue$

$$HR_r = HR_s \cup HR_r \quad (3)$$

Where  $HR_s$  represent the simple heart rate features at ROIs and  $HR_r$  is the ratios of heart rate features. Union of these HR features generate our heart rate features.

### B. Features Standardization & Segmentation

Both the landmark and our proposed heart rate features are on different scales. To fuse the features, we standardized features by partially learning the distribution of features during data loading. We apply standardization as shown in Eq. (4), based on learned distribution over all features.

$$z = \frac{x - \mu}{\sigma} \quad (4)$$

Where  $\mu$  is mean and  $\sigma$  is standard deviation of a feature column.

Our solution works at both the frame and segment level. For segment level operation, we created segments with a length of 30 frames with an overlapping of 10 frames. In our case, the video frame rate is 30 frames per second.

### C. Classification

For the classification task, we need a classifier that should be lightweight and can generalize easily to the new datasets. Classification process should be interpretable so we can follow a directed path for further improvements. To incorporate those requirements, we employed the extreme gradient boosting

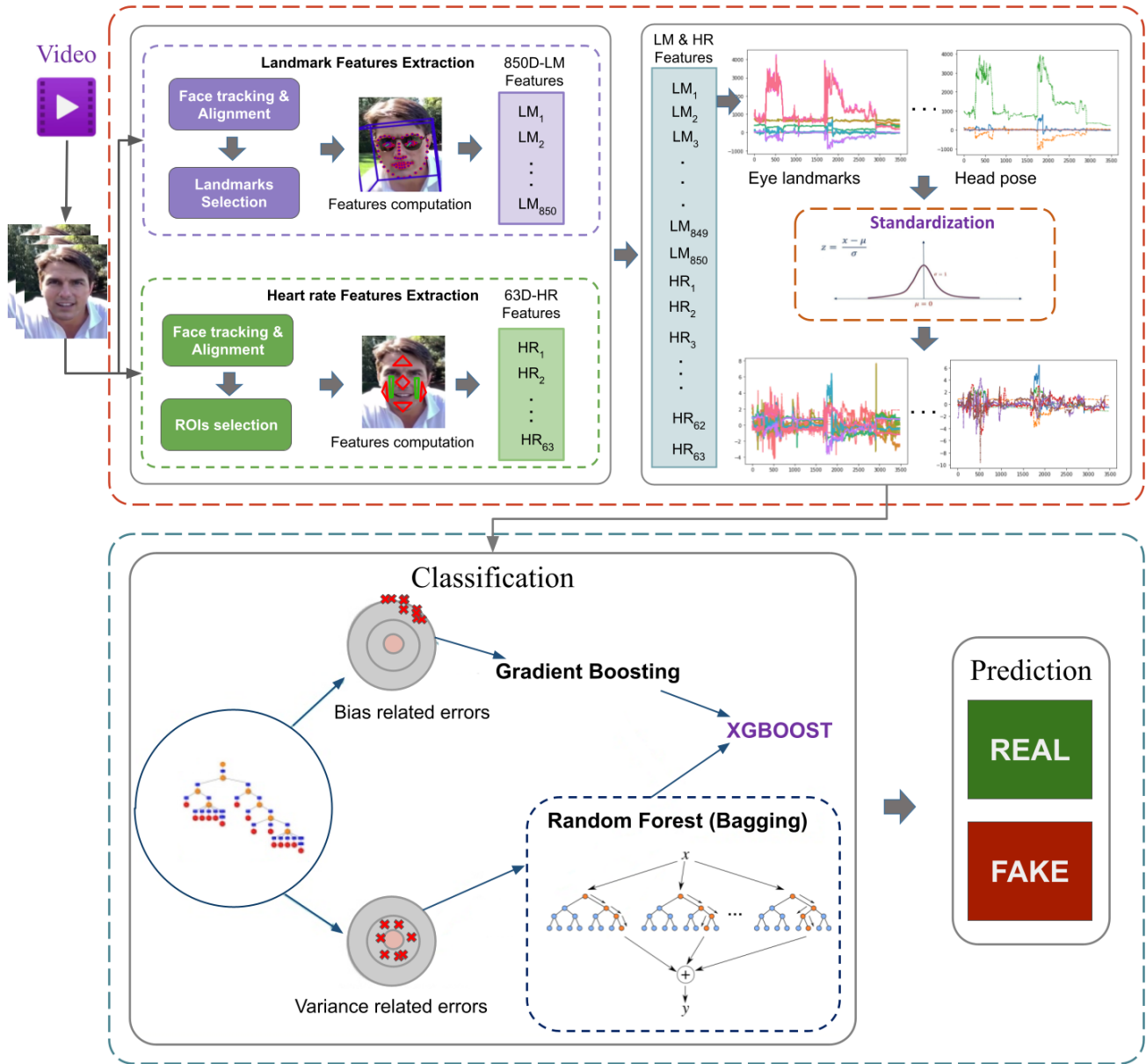


Figure 1. Architecture of the Proposed Framework

(XGBoost) [21], an approach for gradient boosted decision trees. XGBoost is an algorithm in the class of gradient boosting machines. In boosting algorithms, many weak learners are ensemble sequentially to create a strong learner having low variance and high accuracy. In boosting, learning of the next predictor is improved to avoid repeating the error caused by any previous predictor. In Random forest, a model with deeper trees gives good performance but in XGBoost, shallow trees perform better because of boosting. There are two boosting approaches, Adaptive Boosting and Gradient Boosting. Adaptive boosting puts more weight on misclassified data samples. While gradient boosting identifies misclassified samples as gradients using the Gradient Descent to iteratively optimize the loss. XGBoost employs Gradient boosting. Using XGBoost

will be highly effective for large datasets as it is highly scalable and computationally efficient. We can use the power of GPU as XGBoost can perform out-of-core computations. Objective function of XGBoost is based on training loss and a regularization function as shown in Eqs. (5) & (6). Training loss helps in stage wise bagging of trees in the random forest to decrease the variance error. Regularization function helps to reduce the bias related errors using boosting.

$$\mathcal{O} = \sum_{i=1}^n l(y_i, \hat{y}_i^t) + \sum_{i=1}^t \Omega(f_i) \quad (5)$$

Where  $t$  is the total number of trees and  $y_i$  is actual value and  $\hat{y}_i^t$  is the prediction at time  $t$ .  $n$  is the total number of training

samples.

$$\Omega(f_i) = \gamma T + \frac{1}{2} \lambda \sum_{j=1}^T \omega_j^2 \quad (6)$$

Where  $\gamma$  is the min reduction in loss, required for a new split on leaf node and  $\lambda$  is the  $l_2$  regularization term on leaf weights and helps avoid overfitting.

### III. EXPERIMENTS AND RESULTS

#### A. Dataset

We evaluated our method on the world leaders dataset (WLDR) [8]. WLDR is the only dataset with all three types of deepfakes. WLDR comprises real and fake videos of ten U.S. politicians, and real videos of comedy impersonators of those political figures. The WLDR dataset has all three types of deepfakes i.e., face-swap, puppet-master and lip-sync. WLDR has lip-sync deepfakes for only one POI i.e., Obama. Face-swap videos of WLDR are created by replacing the face of the impersonator with the face of the corresponding politician. The WLDR dataset has 1753 real and 93 fake videos. Other datasets like DFDC, FF++ and DFD have more fake videos as compared to real videos. WLDR has more real videos (95%) than the fake videos (5%) which is good, as for better detection, we have to learn the patterns in the real videos rather than fake videos as fake videos are constantly changing with the evolution of GANs. Still it is not large enough to generalize a model to perform well in the wild deepfakes. We used area under the curve (AUC) as an evaluation metric for model evaluation. The reason behind using AUC is that almost all the available datasets are highly class imbalanced. AUC gives a fair performance score for imbalanced classes as compared to Accuracy.

#### B. Performance Evaluation of Proposed Framework

The objective of this experiment is to evaluate the performance of the proposed framework on a diverse dataset WLDR, having all three types of deepfakes. For this purpose, we fed the proposed features of selected landmarks and heart rate features to train the XGBoost based random forest ensemble to perform the classification of bonafide and deepfakes. Heart rate features and sub-categories of landmark features are on different scales. We standardized features before feeding to the classifier. For standardization, we calculated mean and standard deviation of the whole training set during data preparation. We scaled train, test and validation sets to make sure the mean of rescaled data is zero and standard deviation is one. We evaluated our model on frame and segment level. In WLDR, the frame rate of videos is 30 frames per second. For segment level evaluation, we created 30 frames length segments with an overlapping of 10 frames. Our model is robust to both frame and segment level detection.

We evaluated our model on each of six categories of facial landmark and heart rate features. List of features effective to the detection task in descending order is 2D landmark, 3D landmark, eye landmark, headpose, heart rate, shape and action

unit features. Table 1 presents the results of individual feature types. We conducted an evaluation on different combinations of features in the descending order of their effectiveness. Table 2 presents the results of a combination of features categories. We observed from Table 2 that 2D and 3D landmark features are most effective giving an AUC of 0.9311. We also observed that eye landmark and headpose features are effective thereby increasing AUC from 0.9311 to 0.9326, when combined with 2D and 3D landmarks. Additionally, we observed that combining heart rate features with selected landmark features is very effective and increases the AUC from 0.9326 to 0.9505. Based on our observation, we didn't include shape features in the final features-set due to slight improvement in AUC from 0.9505 to 0.9510 when shape features are also included in the fused features-set. Our final features-set includes eye landmarks, headpose, 2D & 3D landmarks and heart rate features. As per our hypothesis, combinations of features that are individually effective also perform better. Finally, we selected five out of seven features categories for our model. We evaluated our model on a wide range of parameters. More specifically, we set the learning rate to 0.01, number of trees to 1500, Max depth tree to 8.

#### C. Performance Comparison of the Proposed and Existing Methods

This experiment is designed to measure the performance of our framework against existing state-of-the-art deepfakes detection methods. For this, we compared the performance of the proposed framework against the [8] and [17]. Table 3 presents the results of comparison of proposed framework against existing models. Our model outperforms [8] that is based on action unit features and derived features capturing mouth movements but our model performance is lower than their extended work [18]. We also compared our model with a deep learning (DL) classifier, LSTM-FCN [22]. Agarwal et al. [8] technique works on the assumption of linear separability of bonafide and deepfake videos in a t-SNE plot based on selected features. But this technique failed to generalize on all types of deepfakes. In their extended work, Agarwal et al. [17] evaluated their method on 10-second video clips rather than frames and segments of small length. Although, this model performs better and generalizes well on all existing datasets of face-swap. However, in this work only face-swap deepfakes are considered and lip-sync and puppet-master deepfakes are not addressed. Moreover, performance of this method [17] is expected to drop if evaluated on frame and segment level due to its threshold based approach. We observed from the results (Table 3) that a DL based model, LSTM-FCN can achieve comparable results as we achieved with XGBoost based Random Forest ensembles. However, compared to LSTM-FCN our proposed framework is light weight and interpretable rather than a black-box oriented model of a DL classifier.

### IV. CONCLUSION AND FUTURE WORK

This work has presented a unified method based on fusion of our novel heart features and facial landmarks for detecting all

TABLE I  
SEGMENT AND FRAME LEVEL AUC ON INDIVIDUAL FEATURES CATEGORIES

Features Used	Eye landmark	Head pose	2D landmark	3D landmark	Shape	Action Unit	Heart Rate
(AUC): Seg level	0.8851	0.8023	0.8982	0.8978	0.7644	0.5027	0.7956
(AUC): Frame level	0.8659	0.7774	0.8903	0.8856	0.7357	0.5017	0.7866

TABLE II  
SEGMENT AND FRAME LEVEL AUC ON COMBINATION FEATURES CATEGORIES

Features Used	2D lmk,3D lmk	Eye lmk,2D lmk,3D lmk	Eye lmk,Headpose,2D lmk,3D lmk	Eye lmk,Head pose,2D lmk,3D lmk,HR	Eye lmk,Head pose,2D lmk,3D lmk,Shape,HR
(AUC): Seg level	0.9297	0.9311	0.9326	0.9505	0.9510
(AUC): Frame level	0.9158	0.9059	0.9068	0.9425	0.9285

TABLE III  
COMPARISON OF XGBOOST WITH [8],[18] AND LSTM-FCN [23]

Model Name	WLDR	Evaluation Levels
Protecting World Leaders [8]	0.93	Frame and segment level
LSTM-FCN	0.95	segment level
XGBoost (proposed)	0.95	Frame and segment level
Appearance and Behavior [18]	0.99	Video level

three types of deepfakes. Unlike many existing methods, our method is light weight, interpretable and effective at the same time. Moreover, compared to existing light weight techniques, our method is more robust and interpretable. We highlighted that an XGBoost based framework is lightweight over the CNN based solutions and generalizes better as compared to other conventional classifiers. For this purpose, we compared our proposed method with a time-series DL classification model, LSTM-FCN. However, proposed framework follows a signature based approach and thus may not be very effective against deepfakes developed in future. Proposed method also need to be enhanced for optimized cross corpus evaluation. For our future work, we'll perform cross-dataset evaluation, experimenting on the datasets that have multiple forgeries per sample.

## V. ACKNOWLEDGMENTS

This work was supported by grant of Punjab HEC of Pakistan via Award No. (PHEC/ARA/PIRCA/20527/21).

## REFERENCES

- [1] (14.06.2021). Reface App. Available: <https://reface.app/>
- [2] (14.06.2021). FaceApp. Available: <https://www.faceapp.com/>
- [3] (14.06.2021) Available: <https://edition.cnn.com/videos/business/2021/03/02/tom-cruise-tiktok-deepfake-orig.cnn-business/video/playlists/business-social-media/>
- [4] Masood, M., Nawaz, M., Malik, K. M., Javed, A., Irtaza, A. (2021). Deepfakes Generation and Detection: State-of-the-art, open challenges, countermeasures, and way forward. arXiv preprint arXiv:2103.00484.
- [5] Agarwal, S., Farid, H., Fried, O., Agrawala, M. (2020). Detecting deepfake videos from phoneme-viseme mismatches. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (pp. 660-661).
- [6] Matern, F., C. Riess, and M. Stamminger. Exploiting visual artifacts to expose deepfakes and face manipulations. in 2019 IEEE Winter Applications of Computer Vision Workshops (WACVW). 2019. IEEE.
- [7] Li, Y. and S. Lyu, Exposing deepfake videos by detecting face warping artifacts. arXiv preprint arXiv:1811.00656, 2018. 2.
- [8] Agarwal, S., et al. Protecting World Leaders Against Deep Fakes. in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. 2019.
- [9] X. Yang, Y. Li, and S. Lyu, "Exposing deep fakes using inconsistent head poses," in ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2019, pp. 8261-8265: IEEE.
- [10] Güera, D. and E.J. Delp. Deepfake video detection using recurrent neural networks. in 2018 15th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS). 2018. IEEE.
- [11] Afchar, D., et al. Mesonet: a compact facial video forgery detection network. in 2018 IEEE International Workshop on Information Forensics and Security (WIFS). 2018. IEEE.
- [12] Nguyen, H.H., J. Yamagishi, and I. Echizen. Capsule-forensics: Using capsule networks to detect forged images and videos. in ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). 2019. IEEE.
- [13] Sabir, E., et al., Recurrent Convolutional Strategies for Face Manipulation Detection in Videos. Interfaces (GUI), 2019. 3: p. 1.
- [14] Yu, N., L. Davis, and M. Fritz, Learning GAN fingerprints towards Image Attribution. arXiv preprint arXiv:1811.08180, 2019.
- [15] (14.06.2021) Available: <https://www.nist.gov/itl/iad/mig/media-forensics-challenge-2018>
- [16] Baltrusaitis, T., et al. Openface 2.0: Facial behavior analysis toolkit. in 2018 13th IEEE International Conference on Automatic Face Gesture Recognition (FG 2018). 2018. IEEE.
- [17] Agarwal, S., Farid, H., El-Gaaly, T., Lim, S. N. (2020, December). Detecting deep-fake videos from appearance and behavior. In 2020 IEEE International Workshop on Information Forensics and Security (WIFS) (pp. 1-6). IEEE.
- [18] Ciftci, U. A., Demir, I., Yin, L. (2020). Fakecatcher: Detection of synthetic portrait videos using biological signals. IEEE Transactions on Pattern Analysis and Machine Intelligence.
- [19] Bonettini, N., Cannas, E. D., Mandelli, S., Bondi, L., Bestagini, P., Tubaro, S. (2021, January). Video face manipulation detection through ensemble of cnns. In 2020 25th International Conference on Pattern Recognition (ICPR) (pp. 5012-5019). IEEE.
- [20] Montserrat, D. M., Hao, H., Yarlagadda, S. K., Baireddy, S., Shao, R., Horváth, J., ... Delp, E. J. (2020). Deepfakes detection with automatic face weighting. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (pp. 668-669).
- [21] Chen, T., Guestrin, C. (2016, August). Xgboost: A scalable tree boosting system. In Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining (pp. 785-794).
- [22] Karim, F., Majumdar, S., Darabi, H., Harford, S. (2019). Multivariate LSTM-FCNs for time series classification. Neural Networks, 116, 237-245.

# Searching for Aesthetical Values in an Upgraded Informal Neighborhood in Tirana

Edmond Manahasa  
Department of Architecture  
Epoka University  
Tirana, Albania  
emanahasa@epoka.edu.al

Arila Rasha  
Department of Architecture  
Epoka University  
Tirana, Albania  
arasha14@epoka.edu.al

**Abstract**— This research aims to explore the possible aesthetical qualities in an informal settlement called Bathore, which evolved in the post-socialist period in capital city of Albania, Tirana. The informal settlement emerged after the fall of socialist system, due to immigration from northern and eastern Albania because of the of unemployment and economic reasons. Based on those considerable parts of agricultural lands at the periphery of Tirana, were usurped to be transformed in informal neighborhoods, which lacked the needed infrastructure. After almost thirty years from this large socio-urban “tectonic crack” these zones are in continuous process of urban upgrade and integration with the other parts of the city through legalization and infrastructure provision. The study benefits from the theoretical framework of aesthetics, to use its main principles in revealing the aesthetical values of the selected urban fragment with Bathore. Based on that it provides an aesthetical analysis examined into two levels: urban and architectural. The urban aesthetical level of this informal context is analyzed by using pattern analysis revealing the existence of proportional relation between building and parcel components. The study found good ratio in between these components, which provided a comfortable built environment. Furthermore, in urban level is found to be used a quasi-modular building footprint in a considerable amount within the selected fragment. Whereas in architectural level similar principles are explored by analyzing the buildings’ exterior visual quality. The study found the existence of aesthetical values also in architectural level. Apart for the unfinished buildings, the aesthetical values in architectural level are provided by the usage of similar/balanced/repeated volumes in non-street side buildings and in street side landmark buildings by qualitative exterior materials and proportional volumes.

**Keywords**—aesthetical values, informal settlements, post-socialist period, Tirana

## I. Introduction on Informal Settlements in Albania

The fall of the socialist regime in Albania is associated with deep consequences in the life of Albanians. Apart from the freedom of speech, the transition to the new liberal economic system caused the bankruptcy of most state companies. As a matter of fact, large number of citizens suffered unemployment, which caused migration to neighbor countries like Italy and Greece. Parallel to this process, immigration also occurred from country’s north and eastern parts to central Albania cities like Tirana and Durres. According Dervishi [1] by the 90s Tirana had 240000 citizens, whereas now it has quadrupled up to 1 million. The existing housing stock could not provide dwelling for this large number of newcomers and, those could not afford to buy a house. This economic reality merged with inefficiencies in urban development and political capital gain reasons according to Fuga [2] made the flourishing of illegal

settlements especially the agricultural lands at the periphery of Tirana.

The housing construction process was based on several steps which were related to the limited budget the settlers possessed. Once that had of the family or one of the males had secured the land plot through different intermediate means (like buying from previous owner or obtaining from the state company who managed during the socialist period; or simply forcefully seizing) the first step was to enclose by a fence or a wall. In the second step after construction of the foundations, sometime only a room or a part of the house was constructed by hand power. At a further step, the first floor of the house was finished, and predominantly there were left bared reinforcement irons at the housings’ terraces aiming to add other upper floors. This process has continued like this and after approximately thirty years these informal dwellings has ended up in three to four floor houses. While some of the informal houses are featured by irregular volumes, in considerable cases due to the qualitative exterior finishing appear like three floor villas.

Because the informal house holders built their homes, without construction permission from 2006 the state institution legalization Agency “ALUIZNI” was established. This process apart from providing a legal status, opened the framework for further infrastructural investments which upgraded the living conditions in the zones. However, the process is continuing and due to the disputable ownership status of the site plot (in which the best case is when it belonged to state and the most complicated one when it belonged to other peoples) have resulted in very long process.

This research aims to reveal the possible aesthetical values of the informal settlements by focusing on a peripheral neighborhood of Tirana, called Bathore. The study selects and analyzes an urban fragment within this neighborhood. Referring to the main principles of aesthetics, it examines the aesthetical values by dividing into urban and architectural levels. For urban level it uses pattern analysis it analyzes the relation between its smaller components, exploring the existence of possible aesthetical principles of scale and proportion. Similarly for architectural level it explores the existence of buildings which possess aesthetical values, by conducting site observation and descriptive analysis of their exterior.

## A. Informal Neighbourhood of Bathore

During the socialist period, Bathore was a village under the administration of Kamza, which was an agricultural town. After the fall of socialist period, as a result on internal migration for reason that were explained in the prior part, newcomers from Northern Albania settled, transforming it

completely in a large an informal zone. Most people obtained the land to develop their dwellings from other settlers who had seized it in the 90s and a minor number simply had occupied the public land [3]. The agricultural character of Bathore have disappeared, to be transformed into a zone into a huge neighborhood fulfilled with informal buildings. Bathore's chaotic silhouette take attention to the eye of any person, who decides to travel through the national road to the northern Albanian cities because the national road passes directly to this area.



Figure I. Location of Bathore in northeast part of Tirana in red dot.

After the establishment of ALUIZNI agency, the process of legalization of informal settlement provides another impetus in offering the needed infrastructure to this settlement. Although due to political intentions, the legalization process have been overextended for almost 30 years, important infrastructural facilities like asphalted roads, solid waste management. Especially the role of one local NGO called CO-PLAN has been essential in producing planning strategies and bridging the gap between the local community and administration through participatory design processes [4]. Apart from that people also has continuously upgraded their houses, improving their dwelling comfort and the exterior quality.



Figure II. The image of first informal settlements in 1994 (left) and the same place in 2007 (middle- ©John Driscoll, IIUD) and the new infrastructure of Bathore in 2019.

## B. Theoretical Unpinning for Urban Aesthetics

The notion of the aesthetics etymologically comes from Greek meaning “to perceive” or related to senses or sensation. Starting from the classical antiquity it has been a subject of elaboration of different field of life like art, music, poetry, or architecture. Plato looked to aesthetics in the absolute beauty, which relied in proportion, harmony, and unity. Aristotle among others emphasized the pleasure as an element of beauty, pointing the importance of perception. Vitruvius provides his vision on buildings proposing his three famous principles: firmitas, utilitas, and venustas, putting the aesthetics the last in order. Especially the influence of Pythagoras’s in this periods perception which used

“mathematical theory of musical consonance” to explain the order of universe is strongly inherent. Other Renaissance masters like Alberti, built their concept on aesthetics by what is “right”, “appropriate”, “proper” and “proportionable” [5]. The Enlightenment period architects put focus on “order” and “sublime”. Together with the emerge of functionalism as the major creative force of modern architecture, indisputably the aesthetical features are criticized to be highly reductionistic. Considering the impact of Loos and Mies Van der Rohe in the 20<sup>th</sup> century apparently rather than decorative features, the spatial qualities have taken the dominant role in architectural compositions.

Proportion and scale in fact are the most fundamental principles of aesthetics widely accepted. While the basis of proportion is laid in the works of Pythagoras, further in later periods till nowadays there are produced different proportional systems like: Golden ratio, Classical orders, Fibonacci numbers, Vitruvian man or Le Modulor, using as its the primary sources nature and human body. As for the scale it refers to how we perceive the size of something in relation to something else. The human scale tends to provide an evaluation of a building size taking as the reference the human body size [6].

Based on this theoretical framework we aim to explore the existence of principles of aesthetics which can be operational in an informal urban context. Apart from the two major principles of aesthetics like scale and proportion, we also aim to see the existence of other principles like symmetry, repetition, composition, balance, linearity, or rhythm. Since the informal settlements are not planned on the contrary are developed spontaneously, we will explore also the possible relation between certain social setting like brotherhood and aesthetical values.

## II. Exploring Aesthetical Values in an Informal Settlement

Having explained the research context and certain basic concepts of aesthetics we have decided to analyze the aesthetical features of selecting an urban fragment within Bathore called neighborhood number 1. To do that we propose to analyze the aesthetical features into two levels: urban and architectural. To comprehend the physical elements in urban level we conduct a city image analysis based on Lynch. Furthermore, to explore the possible aesthetical values in urban level, we use a pattern analysis, revealing the relation between building and parcel and evaluating the existence of proportion and scale principles. In architectural level we firstly define the buildings using their relation to the street into two: street side and non-street-side buildings. In additions we tend to explore the existence of buildings with aesthetical values in this upgraded neighborhood and analyze the possible aesthetical elements and principles in their exterior. We use visual documentation and a descriptive form to reveal these features.

### A. Urban Urban Level Aesthetical Analysis

To search the existence of aesthetical principles in urban level we conduct an urban analysis using Lynch’ elements to comprehend the physical components of the selected fragment. Furthermore, we make a pattern analysis.





Figure III. Location of urban fragment of neighborhood no1 within Bathore map (top) and a 3d drawing of the fragment (down).

### A.1. Comprehending the Urban Fragment Using Lynch's "City Image" Analysis

Based on that the paths, the edges, the districts, the nodes, and the landmarks of the zone are revealed. The selected zone paths consist of three longitudinal streets, the middle of which in the last years is transformed into a dual carriageway street with pavement in both sides, which locally is called a "boulevard". The other streets in perpendicular position, are narrower and considerably are dead ended. The linearity can be considered the only aesthetical element that can be associated with paths. While the two longitudinal streets in both provide well defined edges, in the transversal direction is quite impossible to speak about well-defined edges. The districts are depicted by different colors in Figure 3 and their organizational logic is based on blood relations or place of origin of the householders. It unveils the also how brothers or relatives have bought a common land and further they have developed their houses. While the nodes are quite recessive, the most important landmarks are buildings which are feature with public character like the market building which possesses better exterior quality, or other smaller shops. The shops exterior reflects the transitional state of unfinished buildings.

### A.2. Exploration of Aesthetical Values in Urban Level Through Pattern Analysis

As it was explained before to examine in urban level the aesthetical potential a patterns analysis is conducted (Figure 5). We observe the usage of quasi-similar modular housing footprint, in a considerable amount. Predominantly these buildings are designed with four axes and three spaces in between, producing approximate modules whose width varies around 10 meters. While most of the buildings have a similar size, very low number of buildings possess larger

sizes. One of the buildings which is larger in size is the neighborhood market which three floor high (Figure 4, right-down image). Thus, most of the buildings are in human scale.

Similarly, the ratio between the building footprint to parcel varies from 1/2 to more than 1/5 and the majority is 1/3 and over. These ratios between parcel and housing pattern, due to the existence of considerable green spaces, provide a good balance, resulting comfortable outdoor environment.

The parcel and house pattern of the selected urban fragment is featured by similarity and repetition. Although the logic of parcel division and house scale is directly related on brothers belonging to the same family, it appears that the relation between these two elements seems to be balanced. The repetition of similar house patterns in a balanced way creates the possibilities to have quasi harmonious composition.



Figure IV. Paths within the selected fragment (top) and (down) edges & districts.



Figure V. Nodes (top) and (down) the landmarks within the selected fragment.

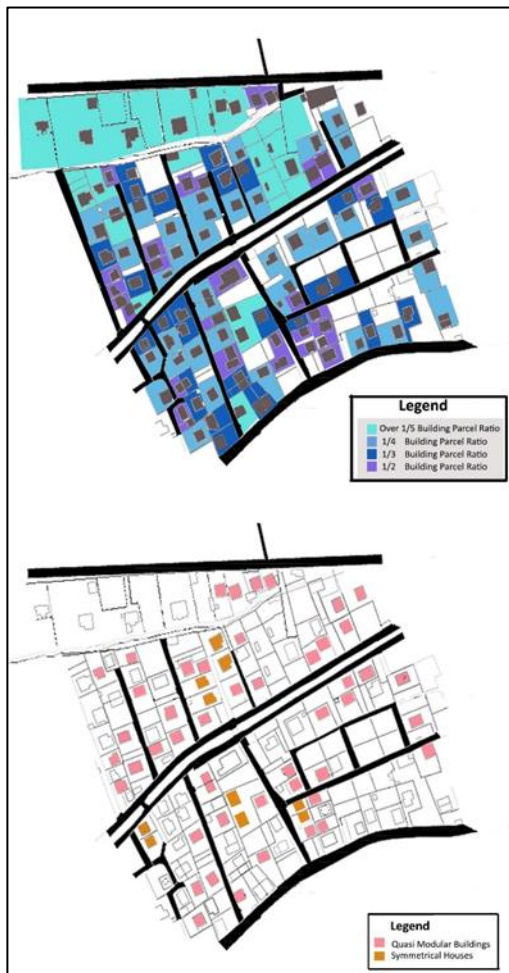


Figure VI. Pattern analysis of parcels and houses relation in selected fragment (top) and usage of quasi-modular building footprint (down).

## B. Architectural Level Aesthetical Analysis

We use two steps to reveal the aesthetical values in architectural level. Firstly, based on an observation we define the building types based on their relation to the street and secondly, we conduct a visual aesthetical analysis in their exterior.

### B.1. Defining The Informal Building Types

The observation on the buildings in the selected urban fragment reveals that they are located into two positions: i. street side and, ii. non-street side. Since there was no street infrastructure, it is quite impossible to say that the selection of the location of those on the streetside has always been done intentionally. However, for the buildings which placed within the surrounding wall of the plot, this can be said that it was done intentionally.

The street-side buildings are observed to be in majority two to three floors high and a small minority are one floor. The one-story building is very rare, due to the commercial values which the presence of the road adds to a space. In this case also most of this typology is used for commercial activity. In the cases of two-story building, in majority of cases the ground floor is used for commercial activities and the first floor for residential functions. In less cases two story buildings both floors are used for commercial purposes. In the case of three-story buildings, they are featured by bigger

dimensions apart from business functions and residential functions, in some cases the offered apartment for renting as also in considerable cases upper residential floors are left empty.



Figure VII. Street-side buildings of selected urban fragment in Demokracia street.



Figure VIII. Non-street side buildings of the selected urban fragment

The non-street side buildings in the selected urban fragment generally are dedicated to housing function and varies from one floor to three floors. The one-story residential buildings have average living conditions. In certain cases, there are "unused" one-story building with a very simple plan, which are constructed only to keep the "occupied land" by instrumentalizing the legalization of that object. The two or more storied houses reflected a good condition as they are built recently or recently restored. However, in majority of cases, their last floors are empty.

### B.2. Exploration of Aesthetical Values Through Visual Analysis of Buildings Exterior

Aiming to find aesthetical potentials of the informal settlements in architectural level, we will make the visual analysis by dividing the building typologies in street-side and non-street side. Apart from the images shown in figures 7 and 8, we observed the existence of upgraded buildings are worthy to be evaluated for their aesthetical values.

The exterior of street side buildings due to the presence of the commercial functions tends to have a more cacophonous visual language which at the same time break the possibility for any rhythmical repetition in the upper floors. This

disbalance is apparent not only in the changing volumes, but also in the varying colors and presence of other additions in the form of staircases, aiming to access the first floor for the same commercial activities. In this context the idea to make the commercial spaces (shops, markets, or other services) more visually appealing, have pushed to cover with heavier materials like marble or stone cladding. This composition generates a visual contrast, between the different functions providing another disbalance. Apart from this apparently those buildings due to incomes coming from commercial activities are upgraded with richer exterior quality, play the role of landmarks.



Figure IX. Street side building with more cacofonic visual qualities (top-right), and landmarks (others) upgraded with richer exterior.

As for the non-street buildings, whose function predominantly is housing their exterior is featured by a more balanced composition. In many cases the same functions, are even repeated in three floors, providing symmetry, repetition, and harmony. In this building typologies this is easier because of the similarity in functions between the floors. In the last years, many buildings have been subject of exterior transformation. This have resulted into villas with vernacular and minor neo-classical architecture elements in exterior, which overpass the prejudice of being an informal house and give possibility to speak about the existence of proportions between the façade elements. An important element that gives values to this kind of house is the existence of abundant front yard, which provides a feeling a pleasure and comfort.



Figure X. Non-street side houses visual qualities in exterior.

Apart from this interestingly, in this small urban fragment there are observed at least four twin houses, which reflect a sense of symmetry. In fact, this approach is based on the relation of brotherhood, in which brothers buy a land plot together and develop identical houses. The similarity in the houses is reflected also in the exterior finishing, where in some cases, the houses are left plastered and uncolored, or are whitewashed, or even in the same colors.



Figure XI. Twin houses placed in symmetrical position.

### III. Concluding Remarks

This study aimed to explore the possible aesthetical values in an upgraded informal neighborhood of Tirana called Bathore selecting an urban fragment within the settlement. Based on the proposed theoretical approach the aesthetical values in such context are analyzed into two levels: urban and architectural. The urban level aesthetical values used pattern analysis to reveal the relation building and parcel. Based on that most of the buildings possess a good ratio with their parcel, varying between  $1/3$  up to  $1/5$ , providing considerable green spaces and consequently a comfortable environment. In addition, we revealed that a considerable amount of the buildings is constructed with four axes and three spaces in between corresponding to approximate quasi-similar modules. These quasi-modules when seen in overall, produce repetition and similarity, resulting in a balanced built environment.

Although still many buildings in the selected fragment remain still unfinished, the study revealed that in parallel with upgrade of the infrastructure by public investments, the building owners also have also adjusted their houses, which we found to have aesthetical values in their exteriors. The study found that buildings in this urban fragment can be defined into two: street side and non-street side. Since street side buildings are possess commercial functions in ground floor, they are featured by more dynamic, sometimes disbalanced volumes, but also in the case of important landmarks they are featured by more qualitative materials in their exteriors, which goes beyond the image of an informal building, producing an added visual aesthetical quality to the neighborhood. The non-street side buildings in majority due to the similarity of functions are featured by a balanced exterior composition. Furthermore, they are featured by repeated volumes and in many cases their exterior is featured also by symmetrical façade solutions. In less cases there are also observed the existence of vernacular and minor neo-classical decoration in their exteriors. Finally, the study

revealed at least in four cases identical twin houses, reflecting a sense of symmetry, which is developed based on the will to have identical houses between brothers. This finding demonstrates that although the informal settlements are not planned, certain social settings like brotherhood, due to this sense of having similar houses can be a reason to develop aesthetical values in a such a built environment.

### References

- [1] Manahasa, Edmond, Place attachment as a tool in examining place identity: A multilayered evaluation through housing in Tirana. PhD dissertation, Istanbul Technical University, 2017.
- [2] Personal Communication with Social Scientist and Philosopher Artan Fuga, 2014.
- [3] Pojani, Dorina, From Squatter Settlement to Suburb: The Transformation of Bathore, Albania. *Hous. Stud.* 2013, 28, 805–821. 2013.
- [4] CoPlan, Enabling a better urban governance. CoPlan case in the areas influenced by illegal buildings, Conference: Strengthening Public Information and Participation for an Open Governance in Albania, Tirana, February 20–21, 2003.
- [5] Scruton, Roger, *The aesthetics of architecture*. Princeton, NJ:Princeton University Press, 1979.
- [6] Rasmussen, Steen Eiler, Eve M. Wendt, *Experiencing architecture*, 1962.
- [7] <http://community.dur.ac.uk/geopad/first-impressions-informal-settlement-bathore-albania/>

# Effect of Oxidation Reactor Structure on Operating Parameters and System Performance in a Nitric Acid Production Plant

Oguzhan Erbas

Department of Mechanical Engineering  
Kutahya Dumlupinar University  
Kutahya, Turkey  
oguzhan.eras@dpiu.edu.tr

F. Menekse Ikbal

Department of Mechanical Engineering  
Kutahya Dumlupinar University  
Kutahya, Turkey  
menekseikbal@gmail.com

Ahmet Akbulut

Department of R&D  
Istanbul Gubre Sanayii A.S (IGSAS)  
Kutahya, Turkey  
ahmet.akbulut@igsas.com.tr

**Abstract**—The factors affecting the process efficiency in a plant producing dilute nitric acid were investigated. It has been observed that the parameter that significantly affects the efficiency in the ammonia combustion process is the reactor unit, which is the heart of the system. It has been determined that most of the malfunctions and stoppages in the existing plant are caused by the problems in the old reactor. Therefore, the aging reactor in the facility was revised, and the reactor unit structure was changed. In this study, the efficiency and performance of the system in the old and new reactor after revision were analyzed.

**Keywords**—nitric acid production, ammonia oxidation reactor, ostwald process, waste heat boiler, energy efficiency

## I. Introduction

With the industrial revolution, there has been a significant increase in the world population. While the world population was 1 billion before the industrial revolution, it is now about 8 billion. The rise in energy demand with population growth, globalization, and income and welfare has led to energy efficiency. Energy efficiency measures how energy losses can be prevented without reducing the production quality and process quantity in industrial enterprises.

The sector that comes to the forefront in energy efficiency studies is the industry-manufacturing field. 2020 has been a challenging year for the world. The world has dealt with the virus epidemic (Covid-19). The stagnation nature of the economic activities of the Covid-19 epidemic has also profoundly affected Turkey. To overcome this crisis, it is necessary to achieve maximum output with minimum input, increase the profit margin without increasing sales, and reduce the costs without reducing the product quality. Energy is the highest cost for the business.

Waste resulting from animal production, nitrogen fertilizer used in plant production, diesel fuel in tractors, thermal fuels in housing, greenhouse, and animal shelters with electricity, the agricultural sector also affects the formation of greenhouse gases that cannot be ignored. While assessments of energy use in agriculture often focus directly on energy use, 50 % and more of the total energy use nitrogen fertilizer production should be considered to be related to energy and other indirect energy uses.

Ammonium nitrate is the most commonly used nitrogen fertilizer. Ammonium nitrate is also used as an explosive. It is obtained by neutralizing nitric acid with ammonia. Depending on the operating conditions, the obtained ammonium nitrate solution has a 50-70 % concentration. After drying the resulting concentrated ammonium nitrate solution, a solid fertilizer is formed. Facility capacities can be selected in a

very flexible range according to needs. The nitrogen value of ammonium nitrate fertilizer is 35 % N [1].

In fertilizer production facilities, diverse starting materials are used depending on the type of fertilizer produced, in some cases together or sometimes separately. The primary starting materials for chemical fertilizer production are ammonia, nitric acid, sulfuric acid, and phosphoric acid. Some of these substances are brought to the facility from outside, while others are produced in the same facility. In an exothermic reaction, gaseous ammonia and nitric acid combine to form ammonium nitrate and water. Nitric acid is preheated without reacting, especially when dilute acid is used; preheating is essential. For this process, steam and hot condensate in the advanced stages of the plant can be used [2,3].

The nitric acid absorber is the primary emission source, and continuous emission into the air occurs from its outlet. These emissions are  $\text{NH}_3$ , nitric acid vapor,  $\text{NO}_x$ ,  $\text{NO}_2$ , and  $\text{NO}$ . The waste gas flow rate and the pollutants in its content may differ according to the process used. For example, the pressure and temperature of the combustion medium, the type/structure of the catalyst, its age compared to its lifetime, the choice of burners, etc. Depending on many factors, varying proportions of  $\text{N}_2\text{O}$  (nitrous oxide) may occur with the combustion gases. Water, typically 0.2 % in liquid ammonia, accumulates in the evaporator as the ammonia is evaporated. Some ammonia is also released when cleaning with intermittent bluffs.

Although ammonia leaks are not common in nitric acid production, they can be a serious source of danger if they do occur. Pipelines, transfer equipment, corrosion punctures, etc., leaks that may occur due to reasons should be monitored by businesses. When preparing air/ammonia mixtures, attention is paid to explosion risks. An additional threat posed by nitrous gas ( $\text{N}_2\text{O}$ ) is the possibility that ammonia, which can accumulate in refrigerated zones, may form salt precipitates in the nitrite/nitrate composition. These carry the risk of explosion, so the risk is eliminated by periodically washing the places where they can occur [4].

For energy efficiency, measures should be created to increase energy efficiency, taking into account the dynamics of each enterprise. In this study, productivity-enhancing studies were carried out in a nitric acid production facility that produces chemical fertilizers. In this nitric acid production facility, processes with high energy consumption were determined, and it was aimed to reduce consumption values.

As a solution alternative to minimize consumption, energy efficiency has been focused on. As a first step, the problems that cause malfunctions in the plant are discussed. Machine-induced stops are examined. It has been seen that this old reactor generates 90% of the failures. The ammonia

oxidation reactor process was analyzed, as the main reason for the downtime at this nitric acid production facility was this old reactor. The reactor was renewed as a result of rehabilitation works at the facility. The operating parameters of the old and new reactors were evaluated, and their efficiency was analyzed [5,6].

## II. The Nitric Acid Production Process and the importance of the ammonium oxidation reactor

Nitric acid (HNO<sub>3</sub>) is obtained by catalytic oxidation of ammonia (NH<sub>3</sub>). This process is called the "Ostwald process". The mixture is passed through a platinum-radium catalyst network. Under normal conditions, the reaction in equation (1) occurs more, and only elemental nitrogen is obtained.



For this reason, a catalyst is used to obtain nitrous oxide. The only catalyst used industrially is the platinum-radium catalyst, which contains 5% to 10% radium and is pencil platinum. It is in the form of a fine mesh. Oxygen atoms are absorbed on the platinum surface.

The reaction takes place between the oxygen atoms on the surface and the ammonia molecules. As a result, the ammonia molecule turns into NO [7]. In the reactor, the gas phase is reversible between NH<sub>3</sub> and oxygen, and an exothermic reaction occurs, and nitrogen oxides (NO<sub>x</sub>) are released (equation 2).



Using 100% HNO<sub>3</sub>, 56 % dilute nitric acid is produced. The raw materials used in the production of nitrate acid are air, water, and ammonia. Nitric acid production steps are as follows;

1. Liquid ammonia is first gasified with water in the ammonia gasifier and then comes to the ammonia superheaters to be heated with hot air. Then the heated gaseous ammonia comes to the ammonia combustion reactor.

2. The air required for the combustion of ammonia is drawn from the atmosphere by the turbocharger and cleaned in the filter. Ammonia with an ammonia concentration of around 10% is mixed with air, passed through the filter, and comes to the ammonia combustion reactor.

3. Nitrous oxide gas (NO) is obtained by burning 11 % ammonia and 89 % air mixture in ammonia combustion plants on platinum-rhodium-palladium catalysts at approximately 870 °C under 2.5–3.5 bar (kg/cm<sup>2</sup>) pressure. Superheated steam is obtained from these hot gases in the waste heat boiler.

4. NO<sub>x</sub> gases coming out of the ammonia combustion reactor at approximately 250 °C are cooled by passing through four coolers. At this stage; It passes first to the rest-gas heat exchanger, then to the boiler feedwater exchanger, and then through the horizontal heat exchanger. Then it reacts by mixing with the secondary air coming from the bleaching column and is cooled again in the vertical heat exchanger. Finally, it enters the oxidation tower last from the bottom.

The oxidation reaction rate of nitrous oxide with additional oxygen proceeds faster at low temperatures. Since the reaction is highly exothermic, severe cooling is required to reach the desired oxidation equilibrium quickly. The nitric acid production process flow chart is given in figure I.

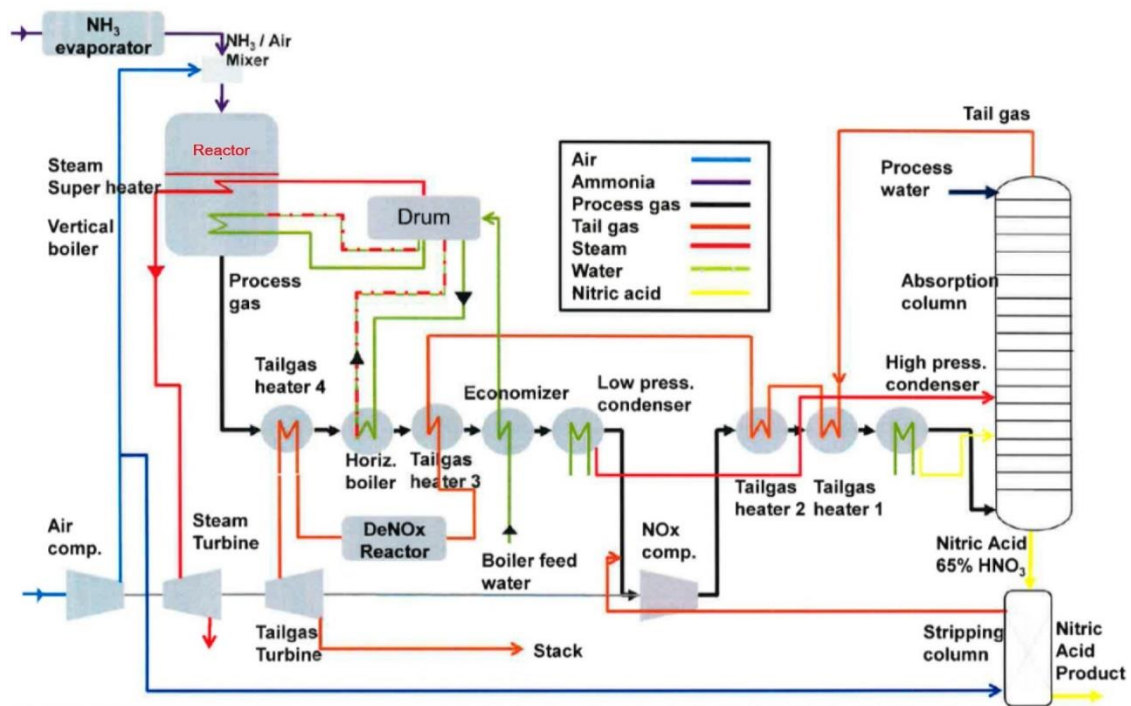


Figure I. Simplified scheme of the Ostwald process for nitric acid production [8]

5.  $\text{NO}_x$  gases entering the tower from the bottom and low-concentration nitric acid ( $\text{HNO}_3$ ) coming from the top react and complete the oxidation. The cooling process in the towers is realized by using the cooling water with the help of the serpentines.

6. Nitrogen oxide gases coming out of the oxidation tower pass to the absorption tower. The water condensed during cooling absorbs nitrogen dioxide from the gas, and acid is obtained at a concentration of around 40% to 50%. This acid accounts for 30% to 40% of the total production. This acid is separated from the gas stream under the cooler, and the absorber is given from a suitable place. Meanwhile, the oxidation state reached around 45%. Additional oxygen from the secondary air helps increase the oxidation state to about 92% to 96%. Nitrous oxide gases coming from the bottom of the tower and 15-16% acid coming from the top (in the absorption column) react and complete its oxidation. Thus, 56% diluted nitric acid is produced [9].

7. Unabsorbable gases leave the tower at a pressure of 2.5 bar and are heated in the rest-gas exchanger. These gases then enter the DeNO<sub>x</sub> system and are discharged into the atmosphere through the chimney. In the DeNO<sub>x</sub> system, with the selective catalytic reduction method,  $\text{NO}_x$  gases are reacted with  $\text{NH}_3$  gas and reduced to  $\text{N}_2$  gas and  $\text{H}_2\text{O}$  vapor already present in the air. As a result, the  $\text{NO}_x$  gas in the flue gas is reduced below the limit value specified in the "Industrial Air Pollution Control Regulation," and the flue gas is analyzed online 24 hours a day, and the data is recorded. In addition, the operating pressure of the system is "medium pressure". Pressure selection is made according to the amount of acid produced (2.5-3.5 bar  $\text{kg}/\text{cm}^2$ )[10].

### III. Results and recommendations

#### A. The main factors affecting process efficiency

When the process in the nitric acid production facility was examined, it was found that the main parameters affecting production efficiency were  $\text{NH}_3$  combustion rate, sub-platinum temperature, no output temperature, amount of steam produced, final amount of acid produced, deactivation times, and postures associated with the oxidation reactor.

The percentage of  $\text{NH}_3$  burning in the mixture is an important parameter. If this  $\text{NH}_3$  ratio increases, there will be an explosion in the system. The rate of ammonia in the ammonia-air mix is kept constant by the automatic regulators (10.5 %  $\text{NH}_3$  + 89.5 %  $\text{O}_2$ ). The facility has recently been automated to minimize human intervention in the process control mechanism. When the  $\text{NH}_3$  rate exceeds 13 %, auto control is activated in the system, and the system is disabled because 15 % is a critical level for  $\text{NH}_3$ .

Another critical parameter is the catalyst (90 % Pt + 10 % Rh). This catalyst is a perforated wire mesh with a diameter of 3800 mm and a wire thickness of 0.06 mm. It is changed every nine months. Its efficiency decreases over time. The conversion efficiency of the ammonia combustion reactor is also an important factor and refers to ammonia consumption. The conversion efficiency decreases with increasing pressure. The structure of the Pt-Rh catalyst is shown in Figure II.

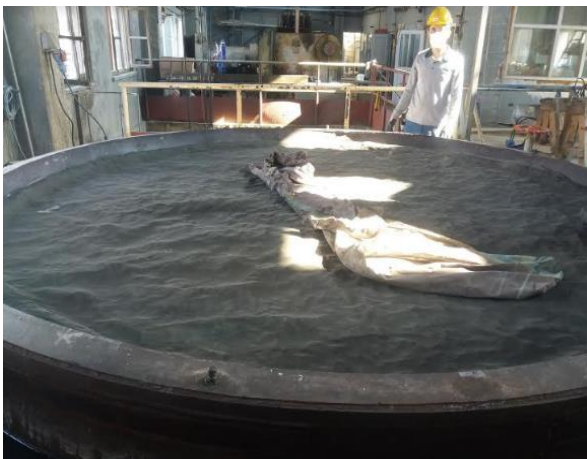


Figure II. Pt-Rh catalyst system structure

The conversion efficiency is 97 % at 870 °C catalyst mesh temperature. The most important feature of this process is that it produces its energy. Once the reaction has started, the pressure and temperature rise should be carefully monitored. Because here sub-platinum temperature is an important parameter.

Since it is worked at high temperatures, stainless steel is used in terms of material strength. In the serpentine pipes, leaks may occur over time, and the system may stop. While there is a loss of time, the production capacity decreases. Therefore, material strength is essential.

The material used is "321(1.4541)" quality stainless steel, which is acid-resistant. To not tire the material in the system, the operating temperatures mustn't exceed the design values.

Another critical parameter is the  $\text{NO}$  outlet temperature. The high  $\text{NO}_x$  gas temperature is coming out of the old reactor before the revision affected the operating parameters negatively. Because the gas temperature should have been 250 °C but found 350 °C. In addition, in the old system, the  $\text{NO}$  gas from the reactor reached 450 °C until it entered the "rest-gas" heat exchanger. This situation has already

reduced the cooling efficiency of nitrous oxide gas entering the heat exchanger at high temperatures. The cooling water in the heat exchangers is also essential because  $\text{NO}_x$  gas affects cooling efficiency. The more cooling is done in the system, the higher the efficiency. Productivity values change according to summer and winter conditions. Production efficiency in summer is almost 10-20 % lower than in winter (atmospheric air is sucked as fluid in the compressor). In body-tube heat exchangers, cracks may occur in the pipes due to high temperatures.

When gas ammonia mixes into the water leaking from the cracks, acid gas is formed in the water that goes to steam and damages the serpentines in the reactor. It erodes the serpentines. The water in the boiler feed exchanger comes to the waste heat boiler in the reactor. For this reason, leaks are not desired in the heat exchangers. When there is a leak, the system stops. Another critical parameter is the pumps. Pumps between the reactor and the high-pressure drum (steam drum) should not be selected with low capacity. If the pumps are insufficient, the circulation cannot be done thoroughly. This can damage the serpentines.

### B. Structure of Ammonia Oxidation Reactor and Effect on System Performance.

The most crucial part of the nitric production facility is the ammonia oxidation reactor. The reactor consists of two parts. These are the combustion part and the steam-generating waste heat boiler part. Catalyst and ammonia burning rates are essential in the combustion part. In the waste heat boiler part, serpentines are imperative. The  $\text{NO}_x$  gas temperature at the reactor exit in the system should not exceed  $250\text{ }^\circ\text{C}$ . However, the design was outdated in the pre-revision system under review. Design efficiency values were very low. The temperature of the reactor exit  $\text{NO}_x$  gas was  $350\text{ }^\circ\text{C}$ . The former reactor and square-shaped serpentines are shown in Figure III.



Figure III. Old reactor and square-shaped serpentines

This high-temperature value was  $450\text{ }^\circ\text{C}$  until it went to the “rest-gas” heat exchanger. The water, which could not be cooled sufficiently in the boiler feed exchanger, was pressed into the waste heat boiler and economizer line. Steam was moving more slowly through the serpentines. As a result of slower progress, cooling did not occur. So it was cracking the material, and it was getting punctured. This was causing leaks in the serpentines, and it was getting punctured. Due to the design of the reactor, there were frequent shutdowns. To prevent this, the reactor was renewed. In the new reactor, the economizer, evaporator, and superheater, which are the waste heat boiler part, were designed in the form of a spiral by placing them on top of each other. Compared to the old one, it became easier to intervene immediately when a malfunction occurs.

Also, the heat transfer surface was almost 10 % more efficient than the other. When the reactor was renewed, there was relief in the heat exchangers. Material life has been increased, and simultaneously, the time has been saved as there are no downtimes as before.

In the old system before the revision, the cause of these critical downtimes is due to leaks and changes. The main problems in this previous system; waste heat boiler gas leakage, casing leakage, preheater (economizer) and evaporator coil leakage, evaporator wall coil leakage, wall pipe leakage, continuous change of casing coils, platinum change and inappropriate  $\text{NO}$  outlet temperature. In figure IV, downtimes by years due to system failures before revision are shown. While there is no stoppage due to malfunction in the new system, it is pretty high in the old system.

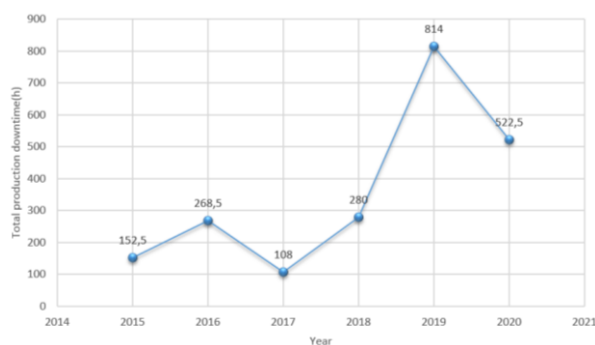


Figure IV. Annual downtime of the facility before overhaul

After the revision, the main innovations made in the reactor are reactor material structure and design dimensions, raching ring basket design, process gas cooler, serpentine structure and surface area, waterways, and automation system. The new reactors after the revision are shown in figure V.



Figure V. New reactors after revision

Many nitric acid furnaces are fixed with a basket filled with Raschig rings to provide structural reinforcement to the sieves, and the partial displacement of Raschig rings determines De- $\text{N}_2\text{O}$  catalysts. Therefore, there is no need to change the basket in general, and the bulk of the sieve mesh is placed as usual. The 50 – 200 mm catalyst layer achieves a high dissolution rate with low additional pressure reduction. However, the pressure drop on the catalyst increases with increasing oxidation pressure.

Old and new reactor rehabilitation studies after revision are evaluated in Table I. After the correction, there was no reactor-induced downtime in the system. In the new system, production stops every nine months for maintenance purposes and the replacement of the sinker.



Table I. Changes in operating parameters before and after revision

Parameter	Before Revision	Post Revision
Sub-platinum temperature	865-870 °C	865-870 °C
NO gas outlet temperature	350-400 °C	250-290 °C
Superheated steam generation	33 t/h	36 t/h
Superheated steam temperature	410 °C	440 °C
Superheated vapor pressure	40 bar	41 bar
Feed water temperature (economizer inlet)	280 °C	150 °C
The final amount of acid produced	610 ton/h	610 ton/h
Deactivation times	90 day	0
Coolant inlet temperature	35 °C	25 °C
Coolant outlet temperature	42 °C	30 °C
Additional water flow	350-400 m <sup>3</sup> /h	50-100 m <sup>3</sup> /h
Reverse current pump operated	2	1
H <sub>2</sub> SO <sub>4</sub> consumption	1 tanker in 10 days	1 tanker in 2,5-3 months
Ammonia flow rate to DeNO <sub>x</sub>	180 kg/h	95 kg/h
Flue gas temperature	185 °C	105 °C
Lamont Differential pressure	No difference	No difference
Dome pressure	40 bar	38 bar
Total compressor airflow	86.500 Nm <sup>3</sup> /h	90.000 Nm <sup>3</sup> /h
Turbine steam inlet flow	No difference	No difference
Amount of Steam produced	20 t/h	26-27 t/h
Ammonia valves	70 % - 80 %	50 % - 60 %
Dome level valve	70 % - 80 %	50 % - 60 %

#### IV. Conclusion

In the dilute nitric acid production facility examined, there was an increase of 3-4 tons per hour in the amount of steam produced with the renewal of the ammonia oxidation reactor. Since the new reactor was commissioned, there has been no reactor-related failure at the facility so far. In the past, it was even possible to exit the circuit twice on the same day. Since there are no downtimes, the production capacity of the plant has increased. In the old reactor, the waste heat boiler part was square, there were problems in the corners of the serpentines turns, and there was more destruction. In addition, there were serpentine leaks due to temperature values. In the new reactor, the serpentines are arranged in a spiral. In the old reactor, the casing cooling pipes were insufficient. That's why there were gas leaks and body punctures. In the new reactor, these problems have disappeared. Since the number of operated reverse current pumps has also decreased, approximately 185 kWh has been saved. As a result; As a result of the new changes, there was no interruption in the production of the plant, and the system efficiency increased by 10%.

#### References

- [1] Hanyu Ma, William F. Schneider, "Structure- and Temperature-Dependence of Pt-Catalyzed Ammonia Oxidation Rates and Selectivities", ACS Catalysis, Volume 9, 2019.
- [2] Jiamin Jin, Ningling Sun, Wende Hu, Haiyang Yuan, Haifeng Wang, Peijun Hu, "Insight into Room-Temperature Catalytic Oxidation of Nitric oxide by Cr<sub>2</sub>O<sub>3</sub>: A DFT Study", ACS Catalysis, Volume 8, 2018.
- [3] Anshumaan Bajpai, Kurt Frey, and William F. Schneider, "Binary Approach to Ternary Cluster Expansions: NO-O-Vacancy System on Pt (111)", The Journal of Physical Chemistry C, Volume 121, 2017.
- [4] Chengxiong Wang, Dezhi Ren, Gavin Harle, Qinggao Qin, Lv Guo, Tingting Zheng, Xuemei Yin, Junchen Du, Yunkun Zhao. "Ammonia removal in selective catalytic oxidation: Influence of catalyst structure on the nitrogen selectivity", Journal of Hazardous Materials, Volume 416, 2021.
- [5] Jiamin Jin, Jianfu Chen, Haifeng Wang, Peijun Hu, "Insight into room-temperature catalytic oxidation of NO by CrO<sub>2</sub>(110): A DFT study", Chinese Chemical Letters, Volume 30, 2019.
- [6] Zhe Hong, Zhong Wang, Xuebing Li, "Catalytic oxidation of nitric oxide (NO) over different catalysts: an overview", Catalysis Science & Technology, issue 16, 2017.
- [7] Ata ul Rauf Salmana, Bjørn Christian Enger, Xavier Auvraya, Rune Lødeng, Mohan Menon, David Waller, Magnus Rønning, "Catalytic oxidation of NO to NO<sub>2</sub> for nitric acid production over a Pt/Al<sub>2</sub>O<sub>3</sub> catalyst", Applied Catalysis A: General, Volume 564, 2018.
- [8] Carlos A. Grande, Kari Anne Andreassen, Jasmina H. Cavka, David Waller, Odd-Arne Lorentsen, Halvor Øien, Hans-Jörg Zander, Stephen Poulston, Sonia García, and Deena Modeshia, "Process Intensification in Nitric Acid Plants by Catalytic Oxidation of Nitric Oxide", Ind. Eng. Chem. Res., Volume 57, 2018.
- [9] Yafei Shen, Xinlei Gea, Mindong Chena, "Catalytic oxidation of nitric oxide (NO) with carbonaceous materials", issue 10, 2016.
- [10] Ruosi Peng, Shujun Li, Xibo Sun, Quanming Ren, Limin Chen, Mingli Fu, Junliang Wu, Daiqi Ye, Size effect of Pt nanoparticles on the catalytic oxidation of toluene over Pt/CeO<sub>2</sub> catalysts, Volume 220, 2018.

# Plant Disease Identification Through Deep Learning

Önsen Toygar

Department of Computer Engineering  
Eastern Mediterranean University  
Famagusta, North Cyprus,  
Mersin, Turkey  
onsen.toygar@emu.edu.tr

Mehtap Köse Ulukök

Department of Computer Engineering  
Bahçeşehir Cyprus University,  
Nicosia, North Cyprus,  
Mersin, Turkey  
mehtap.kose@baucyprus.edu.tr

Emre Özbilge

Department of Computer Engineering  
Cyprus International University,  
Nicosia, North Cyprus,  
Mersin, Turkey  
eozbilge@ciu.edu.tr

**Abstract**—Plant leaves show various symptoms on their surfaces. Image processing and computer vision techniques are applied on leaf images to identify plant diseases. Healthy and diseased plant leaves are involved in several studies to identify or classify disease types using hand-crafted features and deep learning architectures. In this study, we applied a Transfer Learning technique with several deep learning architectures to identify plant diseases from their leaf images. PlantVillage database is used in the experiments with all fourteen plant species and thirty eight classes corresponding to healthy and infected leaf images. Experimental results are presented using several evaluation metrics and a comparison is performed among several deep learning architectures based on Inception, VGG, ResNet, MobileNet, Xception and variants of these architectures. Results are demonstrated in terms of five evaluation metrics, namely accuracy, F1 score, Matthews correlation coefficient, true positive rate and true negative rate. The highest accuracy achieved is 99.81% by ResNet architecture.

**Keywords**—plant disease classification, deep learning, transfer learning, computer vision, image processing

## I. Introduction

Agricultural productivity is so important in countries in which their economy is highly dependent on agriculture. It is stated that 30 to 40% of crops are lost each year through the production chain [1]. Losses from diseases also have an important economic impact, causing a drop in income for crop producers, higher amounts for consumers and distributors. A lot of studies have been carried out under changed environmental conditions, in different locations, to estimate the losses that occur due to different diseases [1]. In this respect, it is vital to detect and identify plant diseases in their initial stage. There exist several computer vision and image processing techniques to identify plant diseases through their leaf images.

Plant-affected diseases emanate from non-living or living factors. The abiotic components known as non-living components include climatically induced situations, contaminated water sources, deficient or surplus liquid for plant use, air and soil chemical substances, nutritional deficiencies, while the biotic components (living) are usually influenced by unwanted weeds, pests, and pathogenic organisms or microorganisms (chromists, fungi, viruses, nematodes, phytoplasm, and bacteria). Therefore, by critically analyzing factors such as ease of spread, ability to exist in diverse forms and complexity in extermination, biotic-borne pathogens continue to present a deleterious threat to the lifespan of plants survival and harvest quality. The three most commonly considered defects for identifying and classifying

plant diseases are spots (caused either by fungi or bacteria), mildew, and rust [2]. Three defect types affecting several plant leaves are shown in Figures 1, 2 and 3. Spot defects are demonstrated on tomato and grape leaves in Figure 1. Rust defects are shown on corn and apple in Figure 2. Mildew defects are presented on cherry and squash leaves in Figure 3.

Every living organism on earth exhibit or react in a particular way when in a condition or situation that deviates from the normal state of being. For example, when the human skin goes red or develops a rash could be due to an allergic reaction or an early indication of an underlying ailment. Plants are not excluded; the leaves in many instances serve as our gateway for diagnosing a lot of diseases in plants. For example, the Early Blight disease of tomato leads to the appearance of small dark spots that expand into circular plaques made up of rings that circumnates on the leaves [3]. This in turn, results in premature defoliation of the leaves and heavy losses in yield. Figure I (a) shows a healthy tomato leaf, and a diseased leaf affected by early blight is shown in Figure I (b).

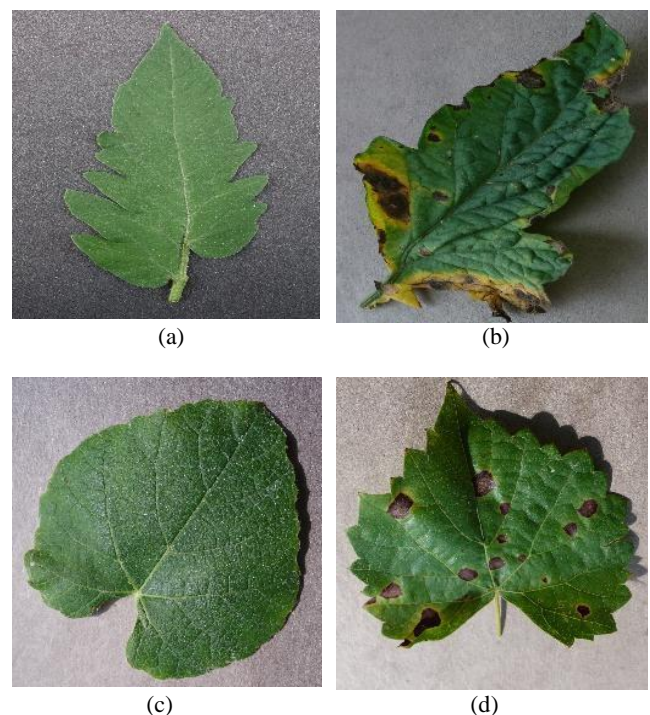


Figure I. Spot defects on Tomato and Grape Leaves ((a) Healthy tomato leaf; (b) Early Blight on tomato leaf; (c) Healthy grape leaf; (d) Black Rot on grape leaf

The fungus, *Guignardia bidwellii* is responsible for the Black Rot disease of grapes. It is usually common in regions of wet, warm and humid climate as this provides a conducive situation for spore germination and infection. This disease spreads when the spores are carried by wind or splashed by rain onto the surfaces of developing plant tissue. This goes on for as long as the environmental conditions remain suitable. Black Rot can be identified when round, tan plaques with dark purple to brown edges are spotted on the leaves. Critical infections may result in leaf deformity, wilting of the leaves [4]. Figure I (c) shows a healthy grape leaf, and a diseased leaf affected by Black Rot is demonstrated in Figure I (d).

On the other hand, *Puccinia Sorghi* Schwein is the pathogen (fungus) responsible for the popular disease of maize known as Common Rust. Early plaques mostly occur in clusters and are circular. But as the plaques ripen, the fungus protrudes through the foliage surface and the plaques elongate with time. The characteristic symptom observed on the maize leaves are Brownish-red oblong pustules, plaques of Common Rust occur on both the upper and lower surfaces of the leaves and are spread sporadically along the leaves. Spores are transported by wind with new infections ensuing weekly or bi-weekly. One plaque is capable of producing both brownish-red urediniospores and black teliospores, yet lastly, only black teliospores will be seen within the plaque [5]. Figure II (a) shows a healthy corn leaf, and a diseased leaf affected by Common Rust is shown in Figure II (b).

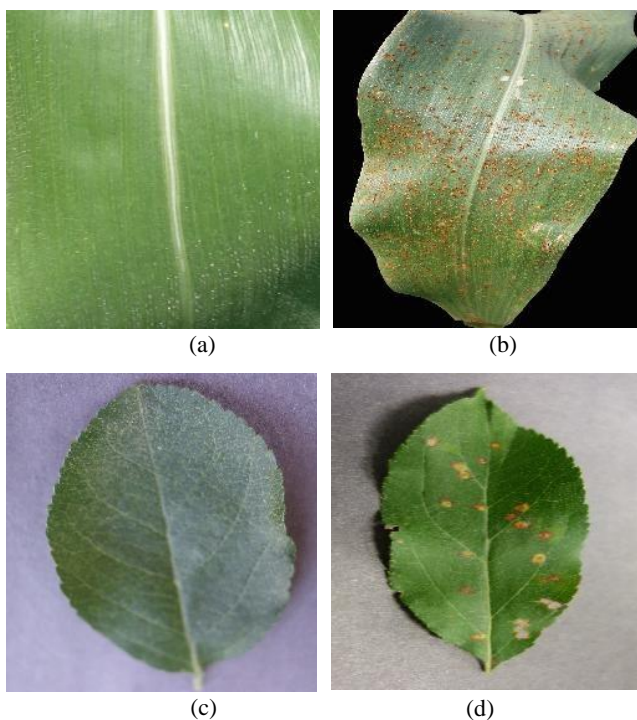


Figure II. Rust defects on Corn and Apple Leaves ((a) Healthy corn leaf; (b) Common Rust on corn leaf; (c) Healthy apple leaf; (d) Cedar Apple Rust on apple leaf

Another sample of rust defects can be seen on Cedar Apple Rust which is a member of the Pucciniaceae family; a class of fungi with several species that typically need two or more host to complete their life cycle. Members of this class are known as rust which are seen at some stage in their evolution and mostly they are orange or reddish in color. The fungus spreads through the leaves and develops aecia beneath the leaves. The aecia produces aeciospores which are wind-blown back to the redcedars. They afterwards germinate and begin gall formation which produce telial horns to restart the process. A heavily infested apple tree can take on a yellowish cast from multiple plaques on the leaves. Figure II (c) shows a healthy apple leaf, and a Cedar Apple Rust infected leaf is presented in Figure II (d).

Moreover, Powdery Mildew is the third type of most common defects seen on some plant leaves such as cotton, cucumber, grape, squash and cherry. Samples of these defects are demonstrated in Figure III on cherry and squash leaves. Figure III (a) shows a healthy cherry leaf, and a diseased cherry leaf affected by Powdery Mildew is shown in Figure III (b). Additionally, two samples of diseased squash leaves affected by Powdery Mildew are presented in Figure III (c) and (d).

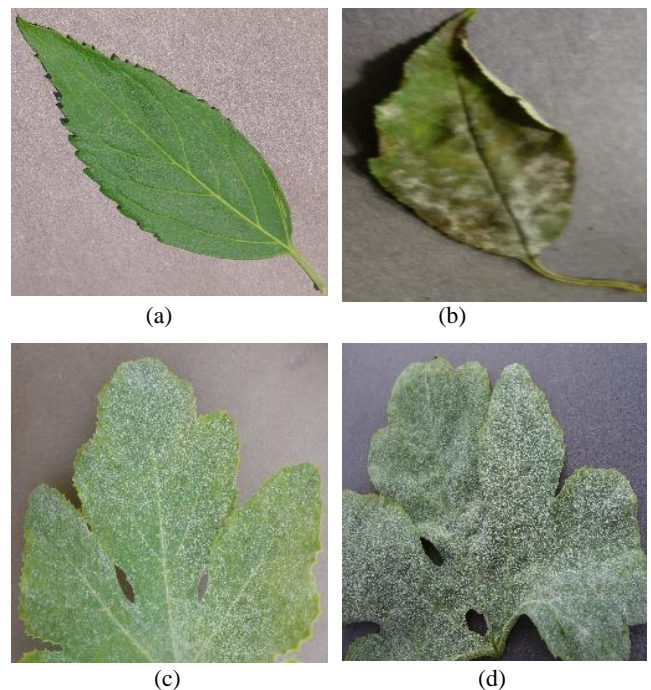


Figure III. Mildew defects on Cherry and Squash Leaves ((a) Healthy cherry leaf; (b) Powdery Mildew on cherry leaf; (c)-(d) Two samples of diseased squash leaves by Powdery Mildew.

The rest of the paper is organized as follows. Literature review is discussed in Section II and the methodology used in this study is explained in Section III. Section IV presents the experiments and results. Finally, Section V concludes the paper with the findings and the summary of the work done in this study.

## II. Literature Review

Plant disease classification has been studied in the literature using two approaches, namely, hand-crafted feature descriptors and deep learning approaches. An extensive research on hand-crafted descriptors for plant disease classification is presented in Kaur et al. [6]. Hand-crafted feature extraction from leaf images requires acquisition, pre-processing, segmentation and feature extraction. Then, these features are used to train classification algorithms such as Support Vector Machines (SVM), Maximum Likelihood Classification (MLC), K-Nearest-Neighbours (KNN), Naive Bayes (NB), Decision Trees (DT), Random Forest (RF) and Artificial Neural Networks (ANN). However, deep learning approaches perform feature extraction and classification automatically.

Deep learning based approaches for plant disease classification are recently presented in several research studies [7]. One of the well-known deep learning techniques is the Convolutional Neural Networks (CNN) and it gains more attention by the researchers. Plant disease classification performance of CNN models is found as superior compared to other classification techniques. However, there are some problems of CNN models with the usage of small datasets. This may cause higher accuracy in classification performance but it is not true in practice. Another drawback is the high execution time that is required with CNN models.

Impact of transfer learning strategies of CNN on pretrained model is worked out by Lee et al. [8]. Plant Village dataset with 38 different classes is studied by using three deep learning architectures, namely, VGG16, InceptionV3, GoogLeNet and a proposed model of GoogLeNetBN. The best accuracy of 99% is achieved with pretrained CNN on that dataset. It is concluded that there is no significant performance difference between pretrained or unpretrained model.

Around two decades' studies for plant diseases detection and classification from plant leaf images by using image processing techniques are summarized in [1, 9]. The importance of digital image quality and its difficulty in real life applications is highlighted in that survey paper. Besides many of the classification techniques, neural network technique gives higher accuracy rate, especially for some plants 100% accuracy is reported.

## III. Methodology

Recently, deep learning based approaches are applied on plant leaf images to classify plant diseases. In this study, we applied Transfer Learning approach to classify plant diseases from leaf images of several plants available in PlantVillage database. All classes of PlantVillage dataset are used in the system and totally there are 38 different classes for 14 plant species. The class numbers and the corresponding disease names are shown in Table I. The details related to the training parameters used in the architecture are given in Table II.

The system architecture is shown in Figure IV using a block diagram of the steps followed for plant disease identification. The system receives input plant leaf images of 256x256x3 size. The images are color images of plant leaves and a pretrained CNN is employed on ImageNet dataset. Afterwards, the frozen weights are trained with backpropagation using a fully connected Artificial Neural Network (ANN) to obtain the outputs of the system as plant disease types to identify the healthy or diseased plant leaves.

The system uses well-known deep networks that are implemented on ImageNet database using Inception [10], ResNet[11], VGG [12], MobileNet [13] and Xception [14] architectures. Pretrained ImageNet network's connection weights are frozen which are used as a feature extractor without retraining the weights. After having obtained the new features from the pretrained network, these features are presented to the custom ANN where its connection weights are learned by using backpropagation algorithm.

## IV. Experiments and Results

Experiments are conducted to perform plant disease classification using healthy and diseased leaf images from PlantVillage dataset. The details related to the experimental setup and the obtained results are presented in the following subsections using several deep learning approaches.

### A. Experimental Setup

PlantVillage database [3] is used in conducting the experiments. All plant species from the database are used in the experiments. Totally, 14 different plant species, namely apple, blueberry, cherry, corn, grape, orange, pepper, potato, raspberry, soybean, squash, strawberry and tomato are used

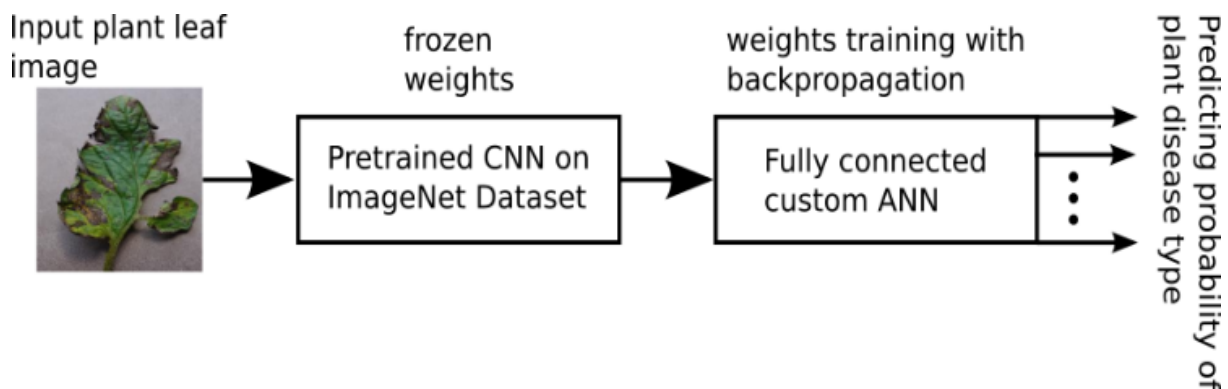


Figure IV. Block diagram of the system architecture with Transfer Learning

with their healthy and/or diseased leaf images. The database consists of 38 different classes that comprise of healthy and different types of infected plant leaves of the aforementioned plant species. The list of classes in PlantVillage database used in the experiments are depicted in Table I. On the other hand, the training parameters used in the experiments are available in Table II including training and test data sizes, input image size, output classes and deep learning architecture's specific parameters.

Table I. List of Classes in PlantVillage Database

1	Apple scab	20	Pepper bell healthy
2	Apple black rot	21	Potato early blight
3	Apple healthy	22	Potato healthy
4	Blueberry healthy	23	Potato late blight
5	Cedar apple rust	24	Raspberry healthy
6	Cherry healthy	25	Soybean healthy
7	Cherry powdery mildew	26	Squash powdery mildew
8	Corn common rust	27	Strawberry healthy
9	Corn gray leaf spot	28	Strawberry leaf scorch
10	Corn healthy	29	Tomato bacterial spot
11	Corn northern Leaf Blight	30	Tomato early blight
12	Grape black rot	31	Tomato healthy
13	Grape esca Black Measles	32	Tomato late blight
14	Grape healthy	33	Tomato leaf mold
15	Grape leaf blight	34	Tomato mosaic virus
16	Orange haunglongbing	35	Tomato septoria leaf spot
17	Peach bacterial spot	36	Tomato target Spot
18	Peach healthy	37	Tomato two spotted spider mite
19	Pepper bell bacterial spot	38	Tomato yellow leaf curl virus

Table II. Training parameters

Training data size	27150
Test data size	27155
Output Classes	38
Input image size	[256x256x3]
Hidden layer (head)	1
Hidden nodes (head)	40
Learning rate	0.01
Dropout rate	0.5
Batch normalisation	enabled
Activation	ReLU
Optimisation algorithm	Adam
Cost function	Focal Cross-Entropy Loss
Maximum epoch	20
Batch Size	16
Random Rotation (data augmentation)	[-36°, 36°]
Random Contrast (data augmentation)	[-10%, 10%]

## B. Experimental Results and Discussion

The experiments are performed using leaf images of several plant species and the experimental results are presented to show the classification of 38 healthy or diseased

classes from 14 plant species. Transfer Learning is applied using several network architectures namely Xception, VGG16, VGG19, ResNet50, ResNet101, ResNet152, ResNet50V2, ResNET101V2, ResNET152V2, InceptionV3, InceptionResNetV2, MobileNet, MobileNetV2. Evaluation of each network architecture is demonstrated using five evaluation metrics. The accuracy (ACC), F1 score, Matthews correlation coefficient (MCC), true positive rate (TPR) and true negative rate (TNR) are computed in each class and presented in Table III as different evaluation metrics. All accuracies are within the range [99.54% - 99.81%] while F1 scores are computed in the range [91.63% - 96.60%]. On the other hand, MCC values are within the range [91.62% - 96.55%] while TPR and TNR values are in the ranges of [91.96% - 96.64%] and [99.74% - 99.90%], respectively.

Table III. Results for several deep learning architectures

Network Architecture	ACC	F1	MCC	TPR	TNR
Xception	0.9959	0.9252	0.9247	0.9273	0.9979
VGG16	0.9969	0.9452	0.9447	0.9463	0.9983
VGG19	0.9965	0.9393	0.9388	0.9410	0.9981
ResNet50	0.9981	0.9660	0.9655	0.9664	0.9990
ResNet101	0.9976	0.9582	0.9577	0.9586	0.9988
ResNet152	0.9978	0.9624	0.9618	0.9627	0.9990
ResNet50V2	0.9970	0.9458	0.9452	0.9470	0.9984
ResNet101V2	0.9970	0.9479	0.9474	0.9491	0.9984
ResNet152V2	0.9968	0.9426	0.9418	0.9428	0.9983
InceptionV3	0.9954	0.9163	0.9162	0.9196	0.9974
InceptionResNetV2	0.9959	0.9224	0.9221	0.9246	0.9978
MobileNet	0.9980	0.9637	0.9632	0.9645	0.9989
MobileNetV2	0.9968	0.9429	0.9424	0.9443	0.9982

Performance evaluation of 13 different deep learning models presented in Table III indicates that all network architectures are robust for identifying healthy or diseased plant leaf images since all the results are high in terms of accuracy, F1 score, Matthews correlation coefficient, true positive rate and true negative rate. In general, all results are above 90% which means that these network architectures are robust and applicable for plant disease classification.

The ranges of performance values for all evaluation metrics show that the minimum performance among 13 different network architectures is obtained by InceptionV3 architecture. Although it has the lowest accuracy among the other deep learning architectures, the performance of InceptionV3 architecture is 99.54% in terms of accuracy. However, the highest performance values for all the metrics are achieved by ResNet50 architecture for plant disease identification. Specifically, the highest accuracy achieved by ResNet50 architecture for classifying 38 classes is 99.81% in terms of accuracy. Therefore, all network architectures employed in the experiments are successful in identifying plant diseases from leaf images.

## V. Conclusion

Identification of plant leaf diseases is studied in this paper. Healthy and diseased plant leaf images from PlantVillage

database are used which constitute 14 plant species and 38 classes. In this study, in order to identify healthy or diseased classes, Transfer Learning is employed using several deep learning architectures, namely Xception, VGG16, VGG19, ResNet50, ResNet101, ResNet152, ResNet50V2, ResNET101V2, ResNET152V2, InceptionV3, InceptionResNetV2, MobileNet, MobileNetV2. Five evaluation metrics are used, namely accuracy, F1 score, Matthews correlation coefficient, true positive rate and true negative rate to compute the performance of each architecture. Experimental results show that all accuracies obtained with the network architectures achieve more than 90% accuracy. The highest performance is achieved by ResNet50 architecture with 99.81% accuracy, 96.60% F1 score, 96.55% Matthews correlation coefficient, 96.64% true positive rate and 99.90% true negative rate. Therefore, ResNet50 is the best architecture among the others used in this study with the highest performance values to identify plant diseases from leaf images.

### References

- [1] Gittaly Dhingra, Vinay Kumar, Hem Dutt Joshi, "Study of digital image processing techniques for leaf disease detection and classification", *Multimedia Tools and Applications*, Volume 77, pp. 19951–20000, 2018.
- [2] Pujari JD, Yakkundimath R, Byadgi AS, "SVM and ANN based classification of plant diseases using feature reduction technique". *Int J Interact Multimed Artif Intell*, Volume 3, pp. 6–14, 2016.
- [3] David P. Hughes and Marcel Salathe, "An open access repository of images on plant health to enable the development of mobile disease diagnostics", *arXiv preprint arXiv:1511.08060*, 2015.
- [4] Angela Madeiras, "Grape IPM- Black Rot", (2019, September 27) Retrieved June 29, 2020, from <https://ag.umass.edu/fruit/fact-sheets/grape-ipm-black-rot>
- [5] Tamra A. Jackson-Ziems, "Rust Disease of Corn in Nebraska", University of Nebraska-Lincoln extension, Institute of Agriculture and Natural Resources. Revised January 2014 Retrieved June 29, 2020, from <http://extensionpublications.unl.edu/assets/pdf/g1680.pdf>
- [6] Sukhvir Kaur, Shreelekha Pandey, Shivani Goel, "Plants Disease Identification and Classification Through Leaf Images: A Survey, *Archives of Computational Methods in Engineering*, 2018.
- [7] M. Nagaraju, Priyanka Chawla, "Systematic review of deep learning techniques in plant disease detection", *Int J Syst Assur Eng Manag* (June 2020), Volume 11, No 3, pp. 547–560, 2020.
- [8] Sue Han Lee, Hervé Goëau, Pierre Bonnet, Alexis Joly, "New perspectives on plant disease characterization based on deep learning", *Computers and Electronics in Agriculture*, Volume 170, (2020) 105220.
- [9] Lawrence C. Ngugi, Moataz Abelwahab, Mohammed Abo-Zahhad, "Recent advances in image processing techniques for automated leaf pest and disease recognition- A Review", *Information Processing in Agriculture*, Volume 8, pp. 27-51, 2021.
- [10] Christian Szegedy, Sergey Ioffe, Vincent Vanhoucke, Alexander A. Alemi, "Inception-V4, inception-ResNet and the Impact of Residual Connections on Learning, Thirty-first AAAI Conference on Artificial Intelligence, 2017.
- [11] Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun, "Deep residual learning for image recognition", *IEEE Conference on Computer Vision and Pattern Recognition*, Las Vegas, NV, USA, 27-30 June 2016.
- [12] Karen Simonyan, Andrew Zisserman, "Very deep convolutional networks for largescale image recognition", *The 3rd International Conference on Learning Representations (ICLR2015)*, 2015.
- [13] Mark Sandler, Andrew Howard, Menglong Zhu, Andrey Zhmoginov, Liang-Chieh Chen, "MobileNetV2: Inverted residuals and linear bottlenecks, *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, UT, USA, 18-23 June 2018.
- [14] François Chollet, "Xception: Deep Learning with Depthwise Separable Convolutions", *IEEE Conference on Computer Vision and Pattern Recognition*, Honolulu, HI, USA, 21-26 July 2017.

# Vaccines Perspective in the COVID-19 Era: Analysis of Twitter Data

Abdulkadir Sahiner<sup>1,2</sup>

<sup>1</sup>Department of Computer Engineering  
Istanbul Sabahattin Zaim University

<sup>2</sup>Department of Mathematical  
Engineering

Yildiz Technical University  
Istanbul, Turkey

Kaan Kemal Polat

Department of Mathematical  
Engineering

Yildiz Technical University  
Istanbul, Turkey

kemalp@yildiz.edu.tr

Hayati Ünsal Özer

Department of Mathematical Engineering  
Yildiz Technical University

Istanbul, Turkey  
huozer@yildiz.edu.tr

**Abstract**— Nowadays, preventive treatments are being developed for the COVID-19 epidemic, which affects all social life. One of these treatments is vaccination studies. Since the subject of vaccination has been discussed by people from past to present, it increases the importance of revealing the approach in this period. In this context, the aim of the study is to determine people's approaches to COVID-19 vaccines through their Tweets. The sentiment analysis method was used within the scope of the study. With this widely used method, the data obtained from Kaggle were analyzed. As a result of the research of the open data set consisting of Tweets related to the Pfizer & BioNTech vaccine, with the LSTM model, in reaching the best accuracy value; The activation function was obtained as "softmax", the epoch value as "10" and the batch size value as "128" appear as remarkable results.

**Keywords**—Vaccine, COVID-19, Twitter, sentiment analysis

## I. Introduction

The COVID-19 epidemic, which affects almost every field, continues its impact today. Vaccine studies have accelerated to prevent COVID-19, which we can describe as a dangerous epidemic that causes the death of people. Today, the vaccines produced by some companies have been taken by countries after the approvals and started to be applied to their citizens.

Vaccine studies regarding COVID-19, which was declared a global epidemic by the World Health Organization (WHO) in March 2020, have been one of the important topics of discussion until today. It is stated that at least 70% of the population should be vaccinated as a precaution with continuity and protection of vaccines [1, 2]. It becomes very important to understand the extent of public support for this situation and to direct society to vaccination accordingly. One of the areas where the public expresses their approach to vaccination is social media platforms [3-5].

Although studies are stating that social media is a very effective tool for vaccination, it is stated that it may contain negative feelings and false information that may affect individual opinions and lead to vaccine rejection [6]. Such a situation has been stated by the World Health Organization (WHO) as one of the ten main factors that are threatening global health [6, 7, and 8]. Due to the differences in the provision of vaccines by countries, people show different approaches to the application of alternative vaccines. Sentiment analyzes are tried to be determined through Tweets on Twitter, one of the social media platforms where people express their opinions against vaccines. Although it is seen that the studies in this context are increasing gradually, some of the studies carried out are as follows.

Bonnevie et al. [9], by examining the Tweets of those who are against vaccines during the COVID-19 period, measuring the anti-vaccination opposition, especially from the beginning of the COVID-19 epidemic in the United States to the present, determining the change in vaccine-related messages and determining the effects of COVID-19 on general vaccine opposition intended. Within the scope of the study, Tweets obtained between 15 February 2020–14 June 2020 were used as data. In the period covering the data set examined in the study, it was stated that the views on anti-vaccination increased by 80%.

A different study, it was aimed to analyze the vaccine discussions on Twitter in the Netherlands. Within the scope of the study, a mixed model was used and a quadruple circular model was followed in which community detection, text-mining, perception analysis, and network analysis were performed. As a result of the study carried out by Kearneys using retweet and igrph packages, it was aimed to contribute to the strategy studies on the vaccine by determining the relational networks of the opposition and shared views on the subject [10].

Balengkship et al. [11] in the study conducted by, it was aimed to determine whether tweets with different emotions and content for vaccination attract different levels of participation (retweets) from Twitter users. As a result of the study using the regression model, it will be the most important step to ensure the participation of key opinion leaders on social media in order to facilitate health education about vaccination in their Tweets and to ensure that their views reach a wider audience. Thus, a positive approach to vaccination will be provided through influential people.

In another study, it was aimed to analyze and determine the data on the international public debate about the pediatric pentavalent vaccine (DTP-HepB-Hib) program by analyzing Twitter messages. In the study in which Twitter data between July 2006 and May 2015 were analyzed, it was seen that there was a little interaction between the tweets, but links containing information about the vaccine were used quite frequently [12].

In the study conducted by Wen-Ying Sylvia Chou & Alexandra Budenz [13], the importance of a data-based communication strategy in controlling the anxiety experienced against vaccine hesitancy was emphasized and in this context, it was aimed to make suggestions regarding the analysis of comments on the vaccine with sentiment analysis on Twitter. Within the scope of the study, it was stated that it was possible to examine the link between the vaccine and the emotional approach and, accordingly, to take precautions against the highly intense opposition to the vaccine with the communication strategy.

Vaccination is one of the most important issues from the past to the present, and today, with widespread communication networks, anti-vaccination is becoming more and more widespread. However, in today's COVID-19 epidemic, where community immunity has a vital stroke, good management of the process and sharing information that will prevent negative opinions against vaccines is possible with data-based methods. In this context, our study is aimed to analyze the shares of people on Twitter about vaccines in the COVID-19 period, with an open data set, using the sentiment analysis method.

## II. Dataset Description

The dataset we used in the study includes the latest tweets about the Pfizer & BioNTech vaccine created by Gabriel Preda on the Kaggle Platform [14]. This data is stated to be collected using the tweepy Python package to access the Twitter API.

In the data set used, there are "id, user\_name, user\_location, user\_description, user\_created, user\_followers, user\_friends, user\_favourites, user\_verified, date, text,

hashtags, source, retweets, favorites, is\_retweet" information about Tweets.

A total of 8631 Tweets were analyzed in the dataset containing the latest tweets about the Pfizer & BioNTech vaccine. Examined Tweets were divided into 3 groups as "positive, negative, and neutral" using Sentiment Intensity Analyzer.

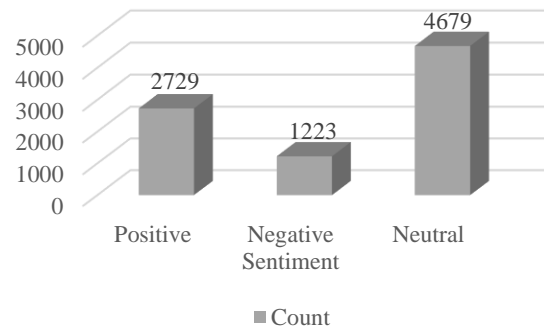


Figure I. Data Sentiment Analysis

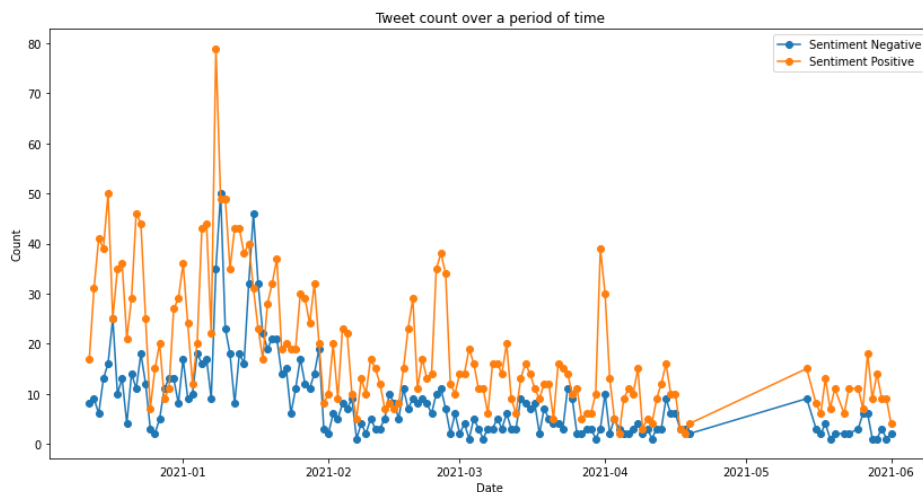


Figure II. Tweet Count Over a Period of Time

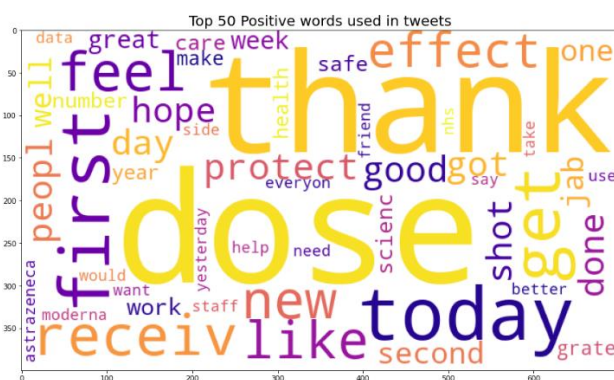


Figure III. Top 50 Positive Words Used in Tweets

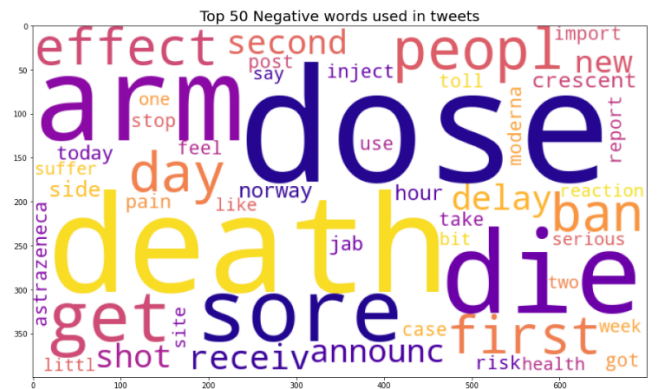


Figure IV. Top 50 Negative Words Used in Tweets

Among the Tweets in the data set used within the scope of the study, words such as "dose, thank" is the most used in the tweets with a positive approach, while words such as "death, arm, die" in the tweets with negative approach attract attention.

## III. Material and Methods

### A. Data Preprocessing

"Word Embedding" was used to process the data and the texts were converted into word vectors. Using the KERAS embedding layer, 32-dimensional word vectors were created.



In addition, various adjustments were made to the data. In this context; hash, internet connections, special characters, single characters, and double spaces are cleared.

Emotion intensity information was added to the cleaned tweets using the VADER [16] sentiment analysis tool.

### B. Long Short Term Memory (LSTM)

Long Short Term Memory is called “LSTM” and is a special type of RNN that can learn long-term dependencies. This model, which was introduced by Hochreiter & Schmidhuber [15], is widely used today due to its effective operation in complex problems.

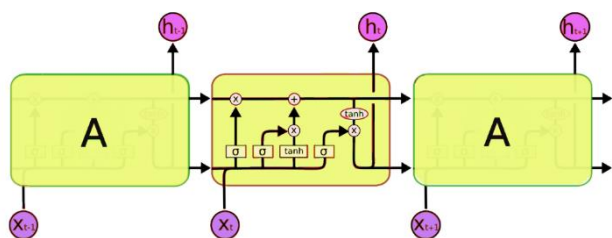


Figure V. Architecture of LSTM Model

One of the tricky issues with natural language processing is that the meanings of words can change depending on their context. In the case of sensitivity analysis, we cannot ignore the occurrence of some words like good because if it comes before the word like not good, its meaning can change completely. This makes it difficult, as it requires reading between the lines. LSTM networks are well suited for solving such problems, as they can remember all the words that lead to what is in question.

The LSTM model used in the study can be summarized as follows:

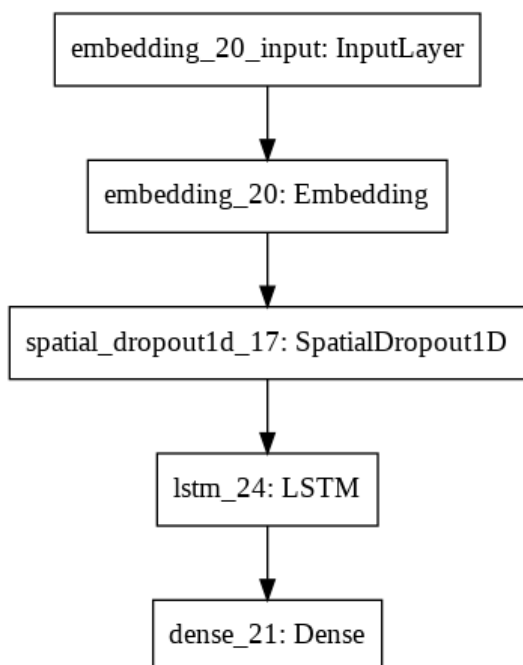


Figure VI. The LSTM Model Used in the Study

### C. Experimental Setup

Python programming language was used in the training of the model proposed in the study. The experiments carried out

within the scope of the study were carried out using the Central Processing Unit (CPU), Graphics Processing Unit (GPU), or Tensor Processing Unit (TPU) hardware and the online cloud service. The parameters used in the experimental process can be summarized in Table I.

Table I. Parameter Settings

Parameter	Value
Activation function	ReLU, softmax
Batch size	128
Loss function	Binary
Optimizer	Adam

Deep learning criteria were used to examine the performance criteria of the model proposed in the study are:

$$Accuracy = \frac{TN + TP}{TN + TP + FN + FP} \quad (1)$$

$$Recall = \frac{TP}{TP + FN} \quad (2)$$

$$F1 - Score = 2 \times \frac{(Precision \times Recall)}{(Precision + Recall)} \quad (3)$$

TP, FP, TN and FN given in Equations (1) - (3) represent the numbers of True Positive, False Positive, True Negative and False Negative, respectively.

In addition, one of the evaluation criteria within the scope of the study is the Matthews correlation coefficient. The Matthews correlation coefficient was introduced by Brian Matthews in 1975 and can be defined as a tool for model evaluation. It is important in terms of revealing the strength of the statistical relationship between the true value and the estimated value [17].

## IV. Results

This section describes the results obtained in this study. As mentioned earlier, one model was used LSTM. More experiments were also performed to investigate different parameters. In this study, optimizers were tested, ADAM optimizer. For model, different parameter values were tested and optimum parameter values were determined. 80% of the data set was used as training data, and 20% was considered as test data. In the created model, 196 cells were used in the LSTM layer.

In the first stage, the results obtained over the different epoch and activation function values applied in the LSTM model are as follows:

Table II. The Effect of Number of Epochs and Activation Function on Parameter Values

Number of Epochs	Activation Function	Accuracy (%)	Recall	F1-Score	Matthews correlation*
10	Softmax	80,83	0,81	0,80	0,66
20	Softmax	79,33	0,79	0,79	0,63
50	Softmax	75,80	0,76	0,75	0,57
10	ReLU	75,51	0,76	0,73	0,55
20	ReLU	79,97	0,80	0,79	0,64
50	ReLU	75,68	0,76	0,75	0,57

\* “The Matthews correlation coefficient is used in machine learning as a measure of the quality of binary and multiclass classifications [18].”

The best performance in the different activation function and epoch values applied in the model used was obtained at the activation function "softmax", and epoch "10" values. In addition, in this case, it was seen that the quality criterion in the model with multiple classifications was quite good (Matthews correlation = 0,66).

In the second stage, the results obtained over the different epoch and batch size values applied in the LSTM model are as follows:

Table III. The Effect of Number of Epochs on Parameter Values\*

Number of Epochs	Batch Size	Accuracy (%)	Recall	F1-Score	Matthews correlation
<b>10</b>	<b>512</b>	<b>80,37</b>	<b>0,80</b>	<b>0,79</b>	<b>0,65</b>
10	128	78,11	0,78	0,78	0,61
10	32	76,38	0,76	0,76	0,58
20	512	79,33	0,79	0,78	0,63
20	128	76,26	0,76	0,76	0,58
20	32	75,91	0,76	0,75	0,57

\* Activation function = softmax

The best performance in the different batch size and epoch values applied in the model used was obtained at the batch size "512", and epoch "10" values. In addition, in this case, it was seen that the quality criterion in the model with multiple classifications was quite good (Matthews correlation = 0,65).

The graph of the change of accuracy and loss value according to epoch value on the data set with the LSTM model is given below.

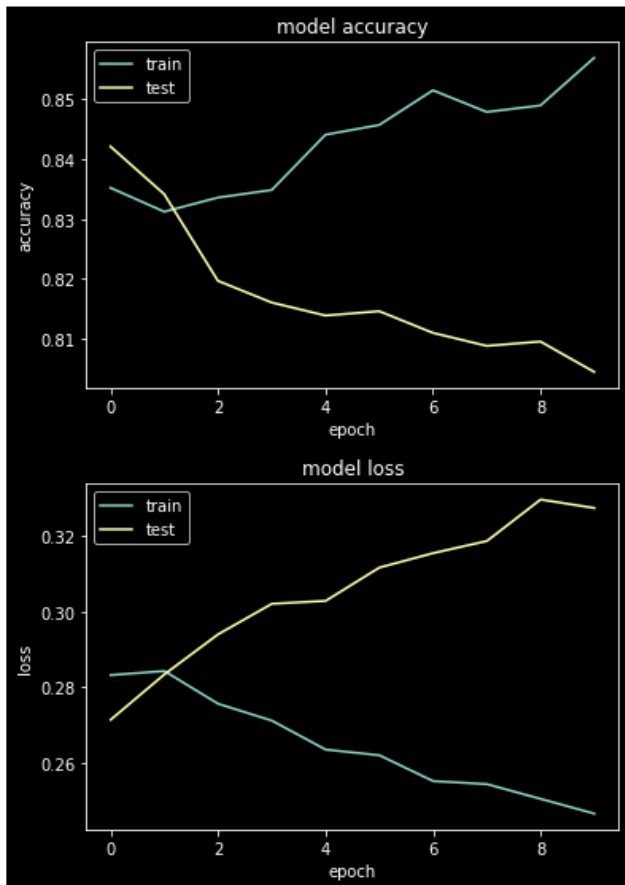


Figure VII. Variation of Model Accuracy and Loss Value by Epoch Value

As a result of the analysis of the open data set consisting of Tweets related to the Pfizer & BioNTech vaccine, with the LSTM model, in reaching the best accuracy value; The activation function was obtained as "softmax", the epoch value as "10" and the batch size value as "128". In addition, the Matthews correlation value reached these values was 0.66.

## V. Conclusion

The main purpose of this study is to analyze the effect of tweets about Pfizer & BioNTech vaccines on anti-vaccine sentiment, to examine how people's views on vaccines have changed during the COVID-19 period. As a result of the use of "ADAM" optimizer and the epoch value of "10", the activation function "softmax" and the batch size value of "128" in the "LSTM" model proposed within the scope of the research, the best accuracy in the data set was obtained with 80.83%.

Although the number of data tagged as positive Tweets is higher than the number of negative Tweets, the fact that the number of neutral Tweets is higher than these two groups shows similar results with recent studies in which anti-vaccination is increasing [19].

Within the scope of the study on how the subject of anti-vaccination, which is an important issue in the literature, differs from the current COVID-19 pandemic, Twitter data was examined within the framework of sentiment analysis. In this process, it is very important for countries to develop strategies according to the approach to vaccines during the pandemic period by examining how user Tweets related to COVID-19 vaccines have changed historically.

Examining the tweets about the vaccine developed by Pfizer & BioNTech within the scope of the study can be expressed as a limitation. In this context, studies can be carried out to examine and compare Tweets related to different COVID-19 vaccines, and as a result, how anti-vaccination differs between these vaccines.

## References

- [1] Orenstein, W. A., & Ahmed, R. (2017). Simply put: Vaccination saves lives.
- [2] Aguas, R., Corder, R. M., King, J. G., Goncalves, G., Ferreira, M. U., & Gomes, M. G. M. (2020). Herd immunity thresholds for SARS-CoV-2 estimated from unfolding epidemics. medRxiv.
- [3] Velasco, E., Agheneza, T., Denecke, K., Kirchner, G., & Eckmanns, T. (2014). Social media and internet-based data in global systems for public health surveillance: a systematic review. *The Milbank Quarterly*, 92(1), 7-33.
- [4] Yousefinaghani, S., Dara, R., Poljak, Z., Bernardo, T. M., & Sharif, S. (2019). The assessment of Twitter's potential for outbreak detection: avian influenza case study. *Scientific reports*, 9(1), 1-17.
- [5] Guess, A. M., Nyhan, B., O'Keeffe, Z., & Reifler, J. (2020). The sources and correlates of exposure to vaccine-related (mis) information online. *Vaccine*, 38(49), 7799-7805.
- [6] Piedrahita-Valdés, H., Piedrahita-Castillo, D., Bermejo-Higuera, J., Guillem-Saiz, P., Bermejo-Higuera, J. R., Guillem-Saiz, J., ... & Machío-Regidor, F. (2021). Vaccine Hesitancy on Social Media: Sentiment Analysis from June 2011 to April 2019. *Vaccines*, 9(1), 28.
- [7] Puri, N., Coomes, E. A., Haghbayan, H., & Gunaratne, K. (2020). Social media and vaccine hesitancy: new updates for the era of COVID-19 and globalized infectious diseases. *Human Vaccines & Immunotherapeutics*, 1-8.

- [8] Kunnean, F., Lambooi, M., Wong, A., Van Den Bosch, A., & Mollema, L. (2020). Monitoring stance towards vaccination in twitter messages. *BMC medical informatics and decision making*, 20(1), 1-14.
- [9] Bonnevie, E., Gallegos-Jeffrey, A., Goldbarg, J., Byrd, B., & Smyser, J. (2020). Quantifying the rise of vaccine opposition on Twitter during the COVID-19 pandemic. *Journal of Communication in Healthcare*, 1-8.
- [10] Lutkenhaus, R. O., Jansz, J., & Bouman, M. P. (2019). Mapping the Dutch vaccination debate on Twitter: identifying communities, narratives, and interactions. *Vaccine: X*, 1, 100019.
- [11] Blankenship, E. B., Goff, M. E., Yin, J., Tse, Z. T. H., Fu, K. W., Liang, H., ... & Fung, I. C. H. (2018). Sentiment, contents, and retweets: a study of two vaccine-related twitter datasets. *The Permanente Journal*, 22.
- [12] Becker, B. F., Larson, H. J., Bonhoeffer, J., Van Mulligen, E. M., Kors, J. A., & Sturkenboom, M. C. (2016). Evaluation of a multinational, multilingual vaccine debate on Twitter. *Vaccine*, 34(50), 6166-6171.
- [13] Chou, W. Y. S., & Budenz, A. (2020). Considering Emotion in COVID-19 vaccine communication: addressing vaccine hesitancy and fostering vaccine confidence. *Health communication*, 35(14), 1718-1722.
- [14] Preda, G. Pfizer Vaccine Tweets. Kaggle Repository: <https://www.kaggle.com/gpreda/pfizer-vaccine-tweets>, 2021.
- [15] Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, 9(8), 1735-1780.
- [16] Hutto, C.J. & Gilbert, E.E. (2014). VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text. Eighth International Conference on Weblogs and Social Media (ICWSM-14). Ann Arbor, MI, June 2014.
- [17] Kaden, M., Hermann, W., & Villmann, T. (2014). Optimization of General Statistical Accuracy Measures for Classification Based on Learning Vector Quantization. In ESANN.
- [18] Scikit-Learn, [sklearn.metrics.matthews\\_corrcoef](https://sklearn.metrics.matthews_corrcoef): [https://scikit-learn.org/stable/modules/generated/sklearn.metrics.matthews\\_corrcoef.html](https://scikit-learn.org/stable/modules/generated/sklearn.metrics.matthews_corrcoef.html), 2021.
- [19] Yousefinaghani, S., Dara, R., Mubareka, S., Papadopoulos, A., & Sharif, S. (2021). An Analysis of COVID-19 Vaccine Sentiments and Opinions on Twitter. *International Journal of Infectious Diseases*.

# Deep Learning-based Healthcare Data Analysis System

Mohammad Ikhsan Zakaria  
Department of Computer Engineering  
Istanbul Sabahattin Zaim University  
Istanbul, Turkey  
zakaria.mohammad@std.izu.edu.tr

Akhtar Jamil  
Department of Computer Engineering  
Istanbul Sabahattin Zaim University  
Istanbul, Turkey  
0000-0002-2592-1039

Alaa Ali Hameed  
Department of Computer Engineering  
Istanbul Sabahattin Zaim University  
Istanbul, Turkey  
0000-0002-8514-9255

**Abstract**— Health care data is playing an essential role in delivering the proper treatment for the patient at the right time. The amount of biomedical data has tremendously increased, which also increased challenges for its analysis. This paper combines supervised and unsupervised learning techniques to obtain insights into the healthcare data for its classification. Specifically, we used a deep neural network to perform anomaly analysis to determine the nature of health problems. Deep convolutional neural network (DCNN) and Deep Autoencoder Convolutional Neural Network (DACNN) were employed to classify the image patterns of extracted electrocardiographs (ECG). We have combined the CNN model with autoencoder to classify heartbeat data. The experimental results showed that high accuracy and lower loss is obtained using our approach.

**Keywords**—Deep Learning, supervised, unsupervised, ECG classification, autoencoder, Convolutional Neural Network

## I. Introduction

Currently health has become the most important thing to people around the world. Most of the countries have tried to counter or prevent the pandemic effect major population by restricted contacts in the community and giving vaccination which they prefer to use. Healthcare system is primary role to handle spreading illness, to care the people. Healthcare data are being analyzed by many scientist to predict what are the coming virus that are being spread in the community. In this paper we propose to make classification in healthcare data system. We focus on heartbeat data classification using deep learning supervised and unsupervised learning.

Heartbeat data can produced by using electrocardiograph (ECG) which is used to record the heart activity in certain time. So basically we see wavelet graph as the results. Most of the time, scientists need to do preprocessing from ECG results raw data. In this paper, we will skip that preprocessing raw ECG files but instead we take prepared data in csv format and analyze them using deep learning algorithms such as Convolutional Neural Network (CNN).

Classification data in healthcare system for heartbeat rate has been analyzed by many studies where they contributed in finding automate system to identify normal beats, supraventricular ectopic beats, ventricular ectopic beats, fusion beats and other abnormal beats[1].

This paper is organized in different sections in the following. Section-I described short explanations about what is the main argument of this paper. Section-II discusses some related works where there are some contributors who have done experiments. Section-III problem statement and section-IV explains about methodology that we used. Section-V discussed some experiments and results. Section-VI discussed about results and conclusion.

## II. Related Works

Class imbalance learning can produce inferior results[2]. Class imbalance can be caused by data sample which come

from different distribution. Moreover, the most standard classifier are designed to optimized the loss function based on how accurate the prediction. However, the predictive accuracy gain low performance metrics from class imbalance learning [3]. As we will see in experiments we found out that heartbeat classes are not balance in distribution. So how to deal with imbalance learning is with preprocessing step where data training is modified to produce more balance class distribution.

The authors [9] described a correlation between sequential forward selection feature selection. Highly correlated features are removed because it will lead to inconsistency models. Sequential forward selection algorithm which is based feature selection can be considered and it is interactive step. The subsets are included to gain the final subset of features where it will obtain correct accuracy.

The authors[3] described that autoencoder to drive deep learning architecture than can learn the hidden representations of data even when data is perturbed by missing values (noises).

Brain Tumor Classification in healthcare systems[6] in order to assist radiologist for a better diagnosis analysis. The classification was investigated using convolutional neural network models by performing extensive experiments using transfer learning with and without augmentation.

## III. Problem Statement

Many attempts to find the best way to achieve higher accuracy to classify highly dimensional and complex ECG data. This research paper shows the gate to find autoencoder system for classifying heartbeat rate category. Meanwhile we keep convolutional neural network be main part of training for supervised and unsupervised learning.

## IV. Methodology

In this paper we propose multi-layer or deep neural network (DNN) in order to find the best classification methods to get the best solution for unsupervised or supervised learning neural networks. DNN can be used for feature extraction and classification of ECG raw data extracted from patient into number of categories.

Our data matrices have high dimensions and imbalance classes makes the training process have become more challenging. So come out with unsampled data to balance classes, the class distribution will be even and good for training or testing convolutional neural networks. Figure 1, is the plotting for heartbeat classification for our four categories which come from ECG signals. This is the sample, so that we know the wavelet visualization.

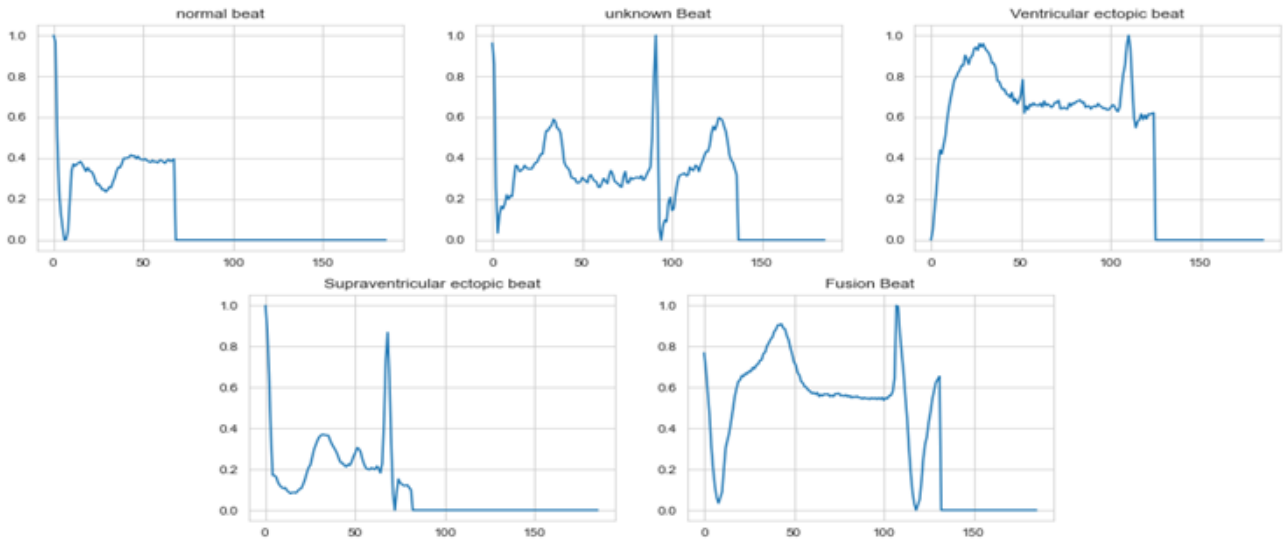


Figure I. Heartbeat Classification

We prefer to use Python programming language because it supports Tensorflow or Keras library and we can code better in Python.

### A. Autoencoder

Basically autoencoders are one of specific feedforward neural networks that use the input same as the output. Inputs are compressed into the lower-dimensional code and then reconstructed the output from this representation. So simply that compression of the input also called the latent-space representation.

An autoencoder can consist of 3 components: encoder, code and decoder. The encoder compresses the input and produces the code, the decoder then reconstructs the input only using this code[3].

## V. Experiments and Results

In this section, we will see some basic requirements for experiments and experiments setup for the test, evaluation of testing for convolutional neural networks such as loss, validation loss and accuracy. The final results obtained ECG data classification experiments and comparative evaluations for several state of the techniques in research field.

### A. Dataset Information

In this research, we have used dataset from “The MIT-BIH Arrhythmia Database” which is available at PhysioNet. There are 48 annotated ECG data, 30 min timeseries long for each excerpts of two-channel ambulatory ECG recordings and obtained from 47 subjects studies by BIH Arrhythmia Lab between 1975 and 1979.

### B. Experiment Setup

Currently, there are several top numerical platform for machine learning development which have many useful features are Tensorflow, Theano, Pandas, Numpy and PyTorch. The platforms are the basic tools which have support library for creating deep learning models. Moreover, Keras Python library can facilitate us to provide the best way to create a range deep learning models on top of Tensorflow.

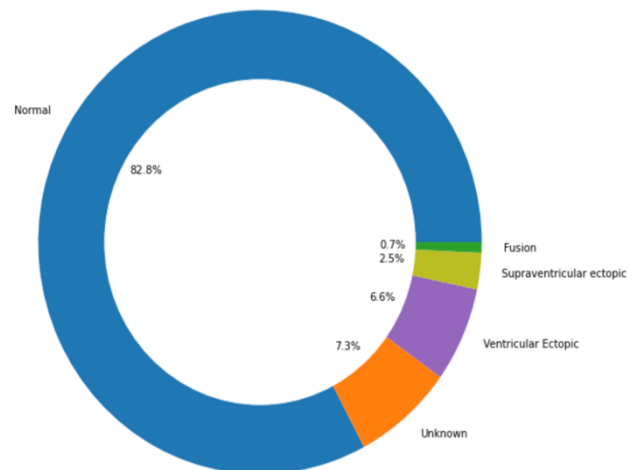


Figure 1 Imbalance Distribution

### C. Data Preprocessing

Before we process the heartbeat data we train using deep learning, we have to check it whether there are missing or empty for each column that represent classes. Moreover if we find imbalance data then we will unsampled and make it evenly distributed. Fig2 shows us that there are four types of heartbeat rate which represent 0.0 for Normal, Supraventricular Ectopic, Ventricular Ectopic, Fusion and Unknown.

### D. Measuring Accuracy and Loss Values

To optimize and measure how strong the deep learning can make prediction, there is always target or result we expect from training. Accuracy implies the number of successful prediction and loss shows how much functions or weight being used but not gaining correct or precise output. Loss output implies how well or bad particular model is behaving after all optimization iterated. Basically we would expect the reduction of loss every time its learning.

Then the test samples are fed to the model and the inaccurate prediction (zero-one loss) the model makes is recorded, after comparison to the true targets. Then the misclassification is calculated. The final measure of accuracy is approaching to 98% using supervised learning. On the other hand using autoencoder we found 78% accurate and has higher loss.

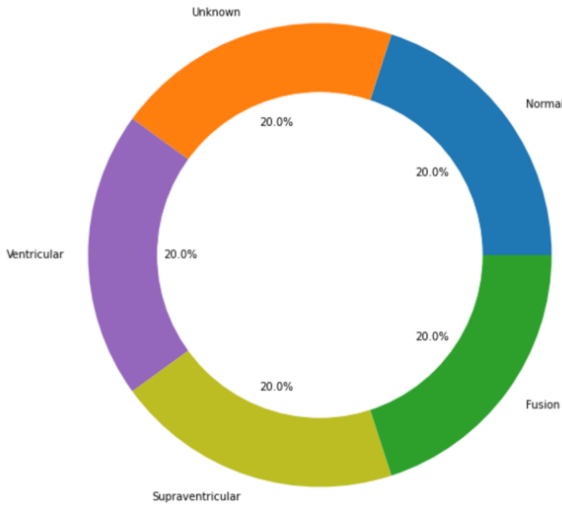


Figure 2 Balance Distribution

Figure 3, shows us plotting data that is already balanced and this could help us to predict or classify heartbeat data using DCNN or DACNN. The shape become more evenly 100000 by 188 matrices. Where column 188<sup>th</sup> is our label for classification.

Following image we will see DCNN architecture that we use in the first stage where we use multilayer convolutional, batch normalization max pooling and with low dense because we need to flatten layer. According to our evaluation DCNN has higher dimension and yielded the best results in moderate time. Larger networks come out with diminishing gradient problems and learning time also significantly longer than other.

In the next level of experimentation, we propose autoencoder to solve healthcare data analysis. As autoencoder detail explained in Fig.5 deep autoencoder neural networks, we provide high input and one line output layer. Because autoencoder supposed to have X\_train as input and output for re-learning.

Figure 4 shows us model accuracy, where training has higher value but validation also gain and almost reaching training accuracy which almost 97%. On the other hand figure 5 shows us model loss, also here training loss lower than validation loss, but it also reaching the most fit which almost 0.025.

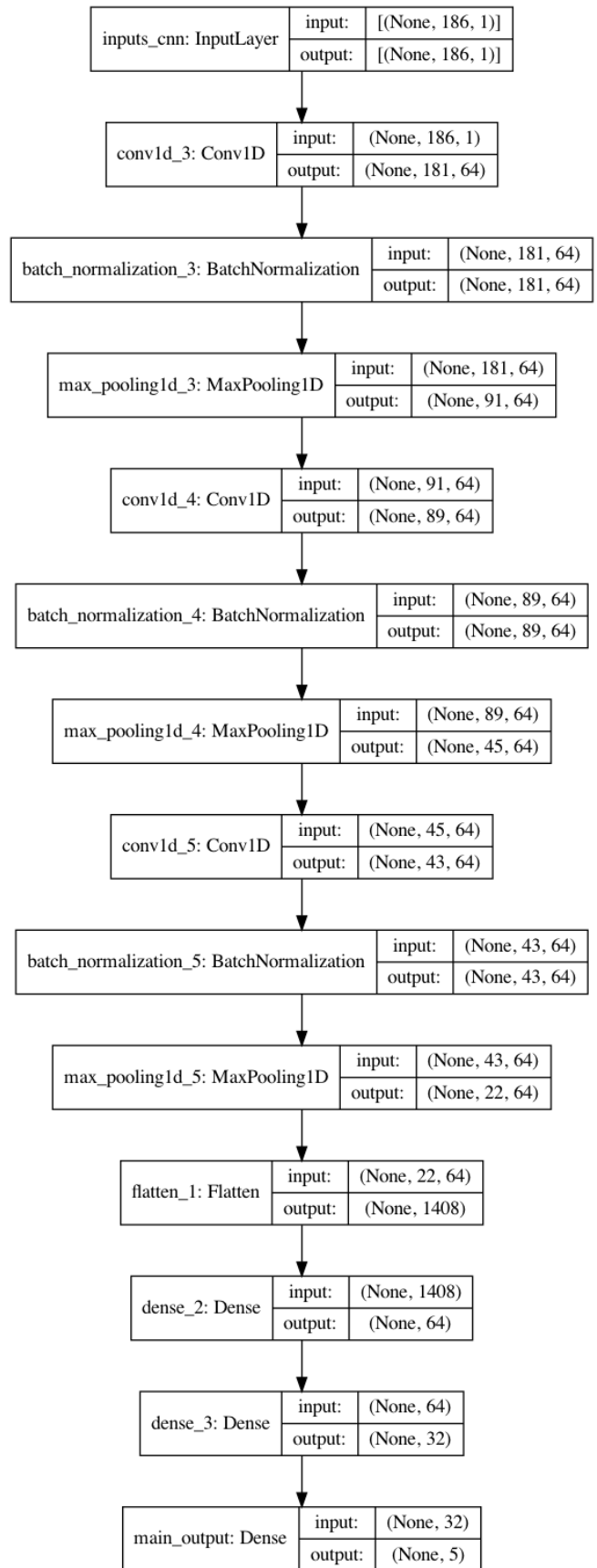


Figure 3 DCCN Architecture

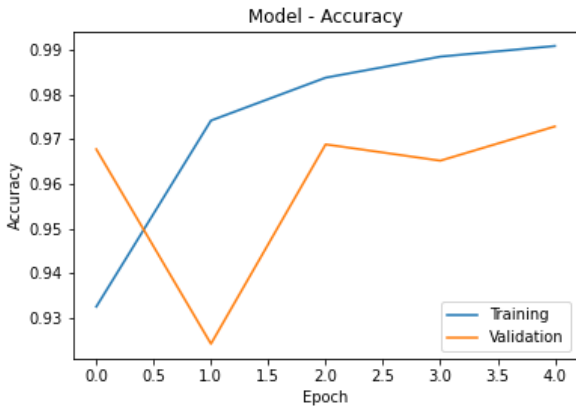


Figure 4 Model Accuracy

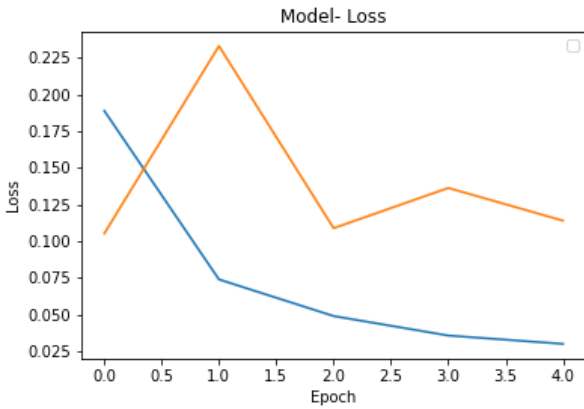


Figure 5 Model Loss

### E. Auto-Encoder

Classification using autoencoder with highly dimension heartbeat data causes a lot of complexity to find the best fit convolutional neural network. So we also make similar preprocessing heartbeat data with balancing and fixing some error during designing models.

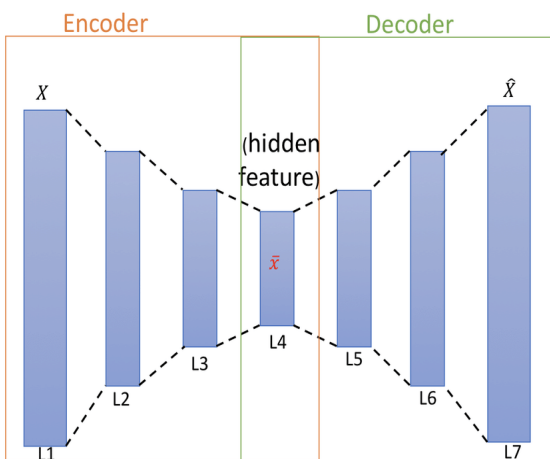


Figure 6 Autoencoder Architecture

Generally autoencoder has four components, which are encoder, decoder, bottleneck and reconstruction loss. Those components cannot be missed in order to get proper or correct models. Encoder means that, the models learn how to reduce the input dimensions into compressed encoded

representation. On the other hand decoder is the reconstruction from encoded representation to be as close as the original input as possible. In this paper we provide Conv1D layer and input (188,1) dimension. After that, we decode using Conv1DTranspose model in order to give flat output and single output. Fig8. Explains details of Deep Autoencoder Convolutional Neural Networks.

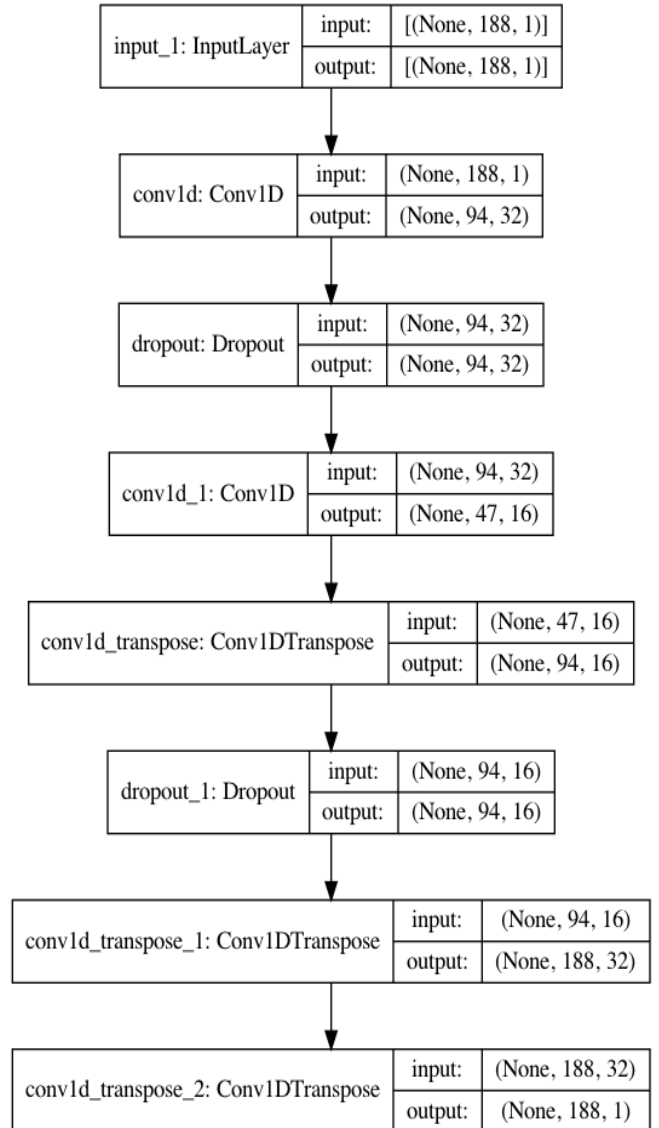


Figure 7 DACNN Architecture

### F. Mean Absolute Error

Machine learning algorithm predictions always come with loss function to minimize the error during training and benefit the implementation. Mean absolute error (MAE) loss will be one of loss that we consider to make evaluation for autoencoder algorithms. We also gave threshold from MAE loss which we can make conclusion again the algorithm that we propose in this paper. Currently we have got 0.0175 threshold from MAE loss which should be best result so far.

Figure 9, explains our calculations again MAE loss in autoencoder architecture.

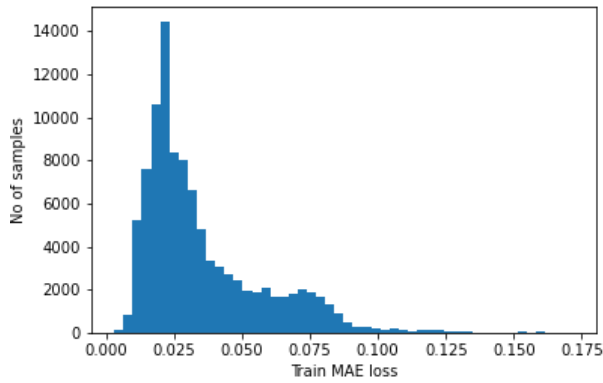


Figure 8 MAE Loss

### G. Prediction

Our important purpose for autoencoder term is that we can make prediction against data we train and automate the prediction. Figure 10, describes the best way on how model convolutional can make better prediction for our large data ECG signals. It shows one sample from Normal beat, which is predicted almost 98% precise.

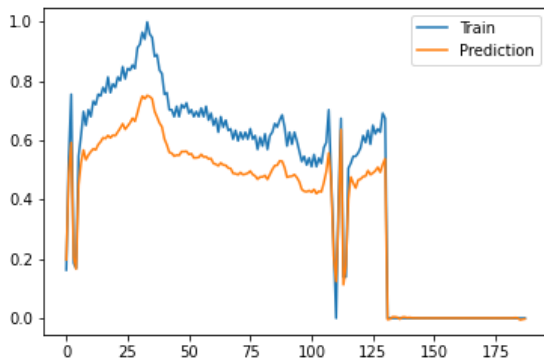


Figure 9 Train and Prediction

### Conclusion

Based on the precision and accuracy, to classify ECG signals model convolutional learning can be combined either as supervised or unsupervised learning. The result can be improved as we modified and re-train neural networks. Sometimes complex data doesn't take more layers for convolutional to make prediction if the target and labeling are

correctly organized. The pattern that we feed to neural network can be easily remember by the machine.

ECG signals data forms from wavelet graph complex data, so when we make prediction using Feed Forward neural network. We can conclude that autoencoder convolutional models can be improved and can be used as common algorithm for Deep learning in healthcare data analysis.

### References

- [1] J. J. Goldberger, M. E. Cain, S. H. Hohnloser et al., "AmericanHeart Association/American College of Cardiology Foundation/Heart Rhythm Society scientific statement on noninvasive risk stratification techniques for identifying patients at risk for sudden cardiac death: a scientific statement from the AmericanHeart Association Council on Clinical Cardiology Committee on Electrocardiography and Arrhythmias and Council on Epidemiology and Prevention," *Circulation*, vol. 118, no. 14, pp. 1497–1518, 2008.
- [2] Muhammad Irfan, Ibrahim A. Hameed, "Deep Learning based Classification for Healthcare Data Analysis System", *IEEE*, 2017.
- [3] Akram Farhadi, David Chen, Rozalina McCoy, Christopher Scott, John A. Miller, Celine M. Vachon, Che Ngufor, "Breast Cancer Classification using Deep Transfer Learning on Structured Healthcare Data", *IEEE, International Conference on Data Science and Advanced Analytics*, 2019.
- [4] Son Phung, Ashnil Kumar, "A Deep Learning Technique for Imputing missing healthcare data", *IEEE*, 2019,
- [5] Xiaokang Zhou, Wei Liang, Kevin I-Kai Wang, Hao Wang, Laurence T. Yang, Qun Jin, "Deep-Learning-Enhanced Human Activity Recognition for Internet of Healthcare Things", *IEEE Journal, IOT*, volume 7, no 7, 2020.
- [6] Khan Muhammad, Salman Khan, Javier Del Ser, Victor Hugo C. de Albuquerque, "Deep Learning for Multigrade Brain Tumor Classification in Smart Healthcare Systems: A Prospective Survey", *IEEE Journal*, volume 32, no 2, 2021.
- [7] Dr. Subiksha.K.P, "Improvement in Analyzing Healthcare Systems using Deep learning Architecture", *ICCCA*, 2018.
- [8] M.M.Al Rahhal, Yakoub Bazi, Haikel AlHichri, Naif Alajlan, Farid Melgani, R.R. Yager, "Deep learning approach for active classification of electrocardiogram signals", *ELSIVIER Journal*, accepted 2016.
- [9] Priyanka Saha, Srabani Patikar, Sarmistha Neogy, "A Correlation - Sequential Forward Selection Based Feature Selection Method for Healthcare Data Analysis", *IEEE Conference*, 2020.
- [10] Xianlong Zeng, Simon Lin, Chang Liu, "Multi-View Deep Learning Framework for Predicting Patient Expenditure in Healthcare", *IEEE Open Journal*, published 2021.
- [11] Pritom Das R, Rakshitha G, Ms. I. Juvanna, Dr. D. Venkat Subramanian, "Retina based Automated Healthcare Framework via Deep Learning", *IEEE Journal, ICGIoT* 2016.
- [12] Tanbin Islam Rohan, Md. Salah Uddin Yusuf, Monira Islam and Shidhartho Roy, "Efficient Approach to Detect Epileptic Seizure using Machine Learning Models for Modern Healthcare System", *IEEE Journal, TENSYP* 2020.



# Optimum Usage of Mixed Damping Systems (Rubber Concrete or X Diagonal Dampers) on Multistory Building

Benni Sami  
 Department of Civil Engineering  
 Gaziantep University  
 Gaziantep, Turkey  
 samibenni@gmail.com

Tayşi Nildem  
 Department of Civil Engineering  
 Gaziantep University  
 Gaziantep, Turkey  
 taysi@gantep.edu.tr

**Abstract**— This research aims to compare usage of X plate diagonal chevron dampers (Win plastic property dampers) on a multi-story building frames in various locations of the structure, accompanied by using treated crumbed rubber concrete with different percentages for the structure's frames. By employing these two systems together as hybrid damping system it will be seen that changing those systems' damping properties will affect the results of push over analysis for the structure such transmitted base shear forces, roof displacement of the performance curve, pseudo acceleration ,pseudo displacement, effective period, ductility ratio and most importantly effective damping ratio which plays major row in reducing demand curve of the design response spectrum. 3D models are modeled and analyzed by ETABS, compare analysis results according to multiple study cases, trying to reach optimum usage of such damping systems according to the indications mentioned above. Where comparison will be focused in this research only on the scale factor of design response spectrum and damping ratio which resulted from push over analysis done for the study cases defined here in this research.

**Keywords**— Damping Systems, Nonlinear Static Analysis (Pushover), Response Spectrum, Earthquakes, Crumbed Rubber Concrete, Chevron Braces

## I. Introduction

In recent years, the rapid construction of skyscrapers has fostered a particular placelessness, as many new skyscrapers have been built without scale, context, or place. By analyzing many skyscrapers and providing numerous visual representations that inspire, stimulate, and engage, Understanding Tall Building asserts that well-designed skyscrapers can rejuvenate cities, spark economic activity, support social life, and enhance city pride. Ground shaking during an earthquake can severely damage structures and the equipment within them. Ground acceleration, velocity, and displacement, when transmitted through a structure, are amplified in most cases.

This amplified motion can produce forces and displacements beyond what the structure can withstand. Many factors affect ground motion and its amplification; therefore, an understanding of how these factors affect the response of structures and equipment is important for safe and economical design. The ground motion of an earthquake is typically measured with strong-motion instruments that record the acceleration of the ground. The recorded acceleration curves, after being corrected for instrument error and baseline, are integrated to obtain the velocity and displacement time histories. The peak ground motion Peak Ground Acceleration (PGA), Peak Ground Velocity (PGV) and Peak Ground Displacement (PGD) are of particular interest for seismic analysis and design. [1]

### A. Resistance and Performance

A seismic performance target has two essential parts - a "damage state" and a "hazard state". Life Safety (LS) and

Immediate Occupancy (IO) are the descriptors of damage states, which are performance objectives only when they relate to a selected seismic hazard level. The hazard may be an earthquake or the probability of ground shaking intensity (10% chance of being exceeded in 50 years). Using the new analysis techniques as a technical tool, it is possible to analyze buildings for multiple performance objectives. Relatively new analysis procedures help to describe the inelastic behavior of the structural components of the building, which allows to estimate the particular behavior of the building during a selected ground motion. The analysis procedure predicts which part of the building will fail first. Because the load and displacement increase, other elements begin to yield and deform inelastically.

The resulting graphical "curve" is a representation of the capacity of the building. Several alternative techniques allow the demand from a selected earthquake or ground shaking intensity to be correlated with the capacity curve to obtain an area on the curve where capacity and demand are equal, which is called the performance point (PP) of the structure, which is an estimate of the particular displacement of the building for the desired ground motion.

### B. Estimation of damping & reduction of 5% damped response spectrum in push over analysis

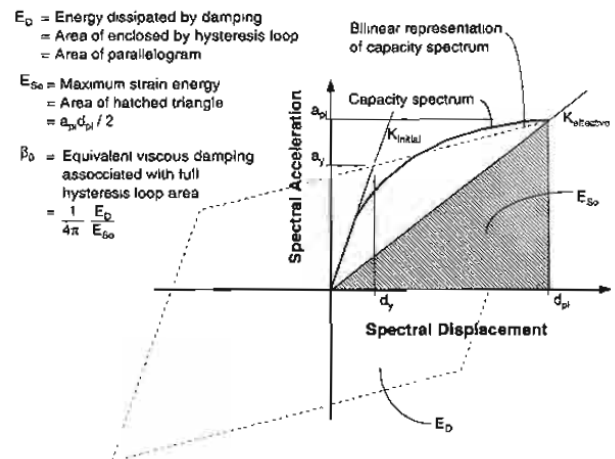


Figure I. Derivation of Damping for Spectral Reduction

The damping that occurs when an earthquake drives a building into the inelastic region can be represented as a combination of viscous damping, which is contained in the structure, and hysteretic damping. Hysteretic damping refers to the area of loops that form when the earthquake force (base shear) is plotted against displacement. It also can be represented as equivalent viscous damping  $\beta_{eq}$ , associated with a maximum displacement of  $d_{pi}$  can be estimated from the following equation:

$$\beta_{eq} = \beta_0 + 0.05 \quad (1)$$

Where,

$\beta_0$  hysteretic damping represented as equivalent viscous damping

0.05 5% viscous damping inherent in the structure (assumed to be constant)

The term  $\beta_0$  can be calculated

$$\beta_0 = \frac{1}{4\pi} \frac{E_D}{E_{S_0}} \quad (2)$$

Where,

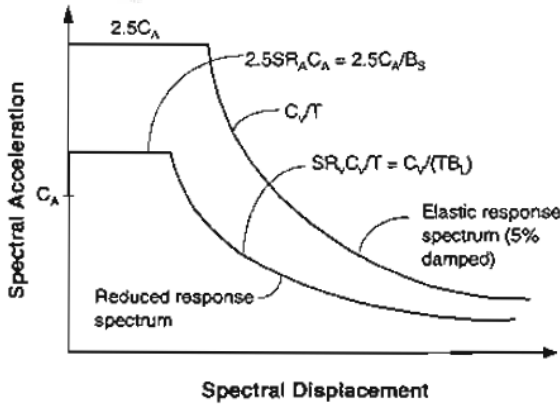


Figure II. Reduced Response Spectrum

$E_D$  the energy dissipated by damping

$E_{S_0}$  the maximum strain energy

$E_D$  is the energy dissipated by the structure in a single cycle of motion, that is, the area enclosed by a single hysteresis loop.  $E_{S_0}$  is the maximum strain energy associated with that cycle of motion, that is, the area of hatched triangle. The term  $E_D$  can be derived as

$$\begin{aligned} E_D &= 4 * (a_{pi} d_{pi} - 2A_1 - 2A_2 - 2A_3) \\ &= 4(a_{pi} d_{pi} \cdot a_y d_y - (d_{pi} - d_y)(a_{pi} - a_y) - 2d_y(a_{pi} - a_y)) \\ &= 4(a_y d_{pi} - d_y a_{pi}) \end{aligned} \quad (3)$$

The term  $E_{S_0}$  can be derived as

$$E_{S_0} = \frac{a_{pi} d_{pi}}{2} \quad (4)$$

$$\beta_0 = \frac{2 a_y d_{pi} - d_y a_{pi}}{\pi a_{pi} d_{pi}} \quad (5)$$

$$\beta_{eq} = \frac{63.6(a_y d_{pi} - d_y a_{pi})}{a_{pi} d_{pi}} + 5 \quad (6)$$

$\beta_{eq}$  values obtained are used to estimate spectral reduction factors, which are used to decrease the elastic (5% damped) response spectrum to reduced response spectrum with damping greater than 5% of critical damping. For damping values less than 25%, spectral reduction factors calculated using  $\beta_{eq}$  equation.

The idealized hysteresis loop is a reasonable approximation for a ductile detailed building subjected to relatively short duration ground shaking (not enough cycles to significantly degrade elements) and with equivalent viscous damping of less than 30%. In order to be consistent with the damping coefficients, B, and to allow simulation of imperfect hysteresis loops (area reduced loops), the concept of effective viscous damping was modified with a damping modification

factor, k, such that the effective viscous damping is,  $\beta_{eff}$  is defined by:

$$\beta_{eff} = k, \beta_0 + 5 = \frac{63.6k(a_y d_{pi} - d_y a_{pi})}{a_{pi} d_{pi}} + 5 \quad (7)$$

k-factor is a measure of the actual structure hysteresis is

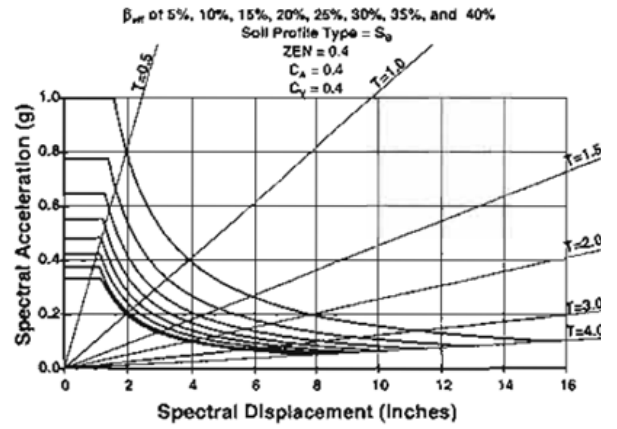


Figure III. Family of Demand Spectra in ADRS Format

well represented, either initially, and/or after degradation. Also, it depends on the structural behavior of the building, and on the quality of the seismic resisting system and the duration of ground motion. There are 3 categories of structural behavior Type A represents stable, reasonably full hysteresis loops ( $k=1$ ). Type B is assigned basic k of 2/3 and represents a moderate reduction of area (k is reduced at higher values of,  $\beta_{eff}$  to be consistent with Type A relationships). Type C represents poor hysteretic behavior with substantial reduction of loop area and is assigned a k of 1/3.

The equations for the reduction factors  $SR_A$  and  $SR_V$  are given by:

$$SR_A = \frac{3.21 - 0.68 \ln(\beta_{eff})}{2.12} = \frac{3.21 - 0.68 \ln \left[ \frac{63.6k(a_y d_{pi} - d_y a_{pi})}{a_{pi} d_{pi}} + 5 \right]}{2.12} \quad (9)$$

$$SR_V = \frac{2.31 - 0.41 \ln(\beta_{eff})}{1.65} = \frac{2.31 - 0.41 \ln \left[ \frac{63.6k(a_y d_{pi} - d_y a_{pi})}{a_{pi} d_{pi}} + 5 \right]}{1.65} \quad (10)$$

## II. Damping systems and effects of adding damping to a structure

Damping is one of many alternative methods proposed for a structure to achieve optimum performance when subjected to seismic, or other types of transient shock and vibration disturbances. The conventional approach would dictate that the structure passively mitigates or dissipate the effects of transient shock through a mixture of strength, flexibility, deformability, and energy absorption. The amount of damping in a conventional structure is incredibly small, and therefore the amount of energy dissipated during transient disturbances is additionally very small. During strong motions, such as earthquakes, conventional structures generally deform far beyond their elastic limits and remain intact only because of their ability to deform inelastically.

Therefore, most of the dissipated energy is absorbed by the structure itself through localized damage. The concept of additional dampers within a structure assumes that some of the

energy added to the structure during a transient is absorbed by additional damping elements rather than the structure itself. An idealized additional damper would be such that the force generated by the damper is large enough and occurs at such a time that the damper forces do not increase the overall stress within the structure. Properly implemented, a perfect damper should be able to simultaneously reduce both stress and deflection within the structure [3].

### A. X-shaped metallic dampers

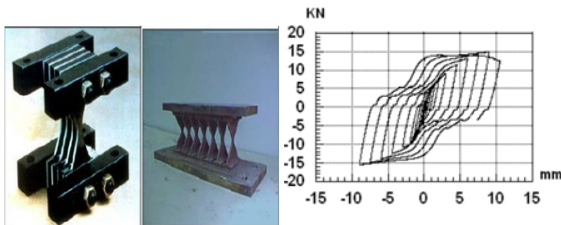


Figure IV. Samples of X Plate Dampers with Wen Hysteretic Loops

X-plate dampers consist of one or more X-shaped steel plates, each plate having a double curvature and arranged in parallel; the number of plates depends on the amount of energy required to be dissipated in the given system. The material used to make the X-plate can be any metal that allows a large amount of deformation, such as mild steel, although sometimes lead or more exotic metal alloys are used. To reduce the response of the structure by dissipating the applied seismic energy, such a damper may be used with a suitable support system, where a combination of braces and XPDs may be used in the building structure and such an assembly is known as a device brace. When such a system is subjected to lateral forces such as earthquakes, high winds, etc., the seismic energy introduced is dissipated by their flexural deformation. They can withstand many cycles of stable yielding deformation, resulting in a high degree of energy dissipation or damping.

The goal behind using X-shaped dampers is to have a constant strain variation over their height to ensure that yielding occurs simultaneously and uniformly over the entire height of the damper. XPDs can also behave nonlinearly, but limit the behavior of the structure to the linear-elastic region. In a series of experimental tests, the behavior of XPDs was studied and the following results were observed: It exhibits

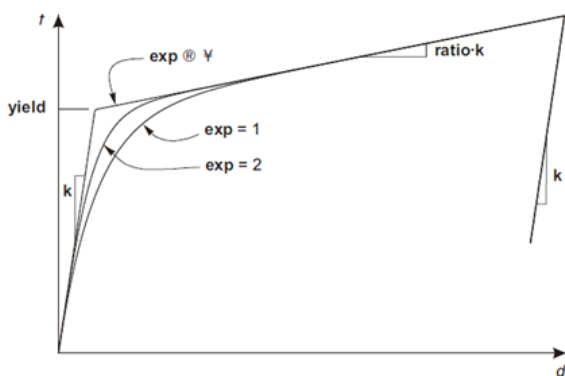


Figure V. Nonlinear Force- Deformation relationship of Wen Plasticity

smooth nonlinear hysteretic loops under plastic deformation; it can withstand a large number of yield reversals; there is no significant stiffness or strength degradation; and it can be accurately modeled by Wen's hysteretic model or as a bilinear elasto-plastic material [4].

(For X Plates Dampers, XPD) for each deformational degree of freedom, it may specify independent uniaxial plasticity properties. The plasticity model is based on the hysteretic behavior proposed by Wen (1976).

All internal deformations are independent. The yielding at one degree of freedom does not affect the behavior of the other deformations. If it does not specify nonlinear properties for a degree of freedom, that degree of freedom is linear using the effective stiffness, which may be zero.

The nonlinear force- deformation relationship is given by:

$$f = \text{ratio } k d + (1 - \text{ratio}) \text{yield } z \quad (11)$$

where  $k$  is the elastic spring constant,  $\text{yield}$  is the yield force,  $\text{ratio}$  is the specified ratio of post- yield stiffness to elastic stiffness ( $k$ ), and  $z$  is an internal hysteretic variable. This variable has a range of  $|z| \leq 1$ , with the yield surface represented by  $|z| = 1$ . The initial value of  $z$  is zero, and it evolves according to the differential equation:

$$\dot{z} = \frac{k}{\text{yield}} \begin{cases} \dot{d}(1 - |z|^{\text{exp}}) & \text{if } \dot{d} z > 0 \\ \dot{d} & \text{otherwise} \end{cases} \quad (12)$$

where  $\text{exp}$  is an exponent greater than or equal to unity. Larger values of this exponent increase the sharpness of yielding. The practical limit for  $\text{exp}$  is about 20. The equation for  $\dot{z}$  is equivalent to Wen's model with  $A=1$  and  $\alpha=\beta=0.5$

### III. Rubber Concrete

Crumb Rubber can absorb sudden shocks by controlling the motion of waves transmitted by moving loads. Rubber compresses and deforms easily, but the rate of deformation decreases as the load increases, making it a good shock absorber. Rubber has very low hysteresis and controls energy dissipation in a system subjected to vibration-induced forces. Rubber has an inherent damping ability and this property can be exploited to improve the damping and vibration characteristics of concrete. In addition, rubber is lightweight and non-corrosive, making it easy to use and apply. It is relatively inexpensive compared to other conventional dampers such as steel springs and tuned mass dampers.

Material scientists have attempted to form concrete with a ductile material. However, it appears that due to the brittle nature of concrete, the most direct and effective approach to creating damage tolerant concrete structures would be to embed intrinsic tensile ductility into concrete. If concrete behaves like steel in tension (highly ductile) while retaining all other advantages (e.g., high, and extreme compressive strength), concrete structures with improved serviceability and safety can be easily realized. In addition, crumb rubber can absorb sudden impacts by controlling the movement of waves transmitted by moving loads. Rubber compresses and deforms easily, but the rate of deformation decreases as the load increases, which makes it a good damper. Rubber has low hysteresis and controls energy dissipation in a system that is highly exposed to vibration-induced forces. Structures located near roads are subject to vibrations caused by moving vehicles. These vibrations are harmful to the pavement and structures adjacent to the pavement. The use of rubber in

pavements as a substitute for natural aggregate can reduce this effect [5].

#### IV. Optimum usage of mixed damping

Nowadays, technological development and the practical solutions it provides to complex issues have become an urgent necessity that cannot be dispensed with. Accordingly, the development of damping systems used in modern structures and high-rise buildings depends on the software provided by computers mainly. Three-dimensional models with study cases were modeled using ETABS to support this research based on the assumptions and equations that were mentioned in the previous chapters.

The main objective of mixing damping systems (Rubber Concrete, chevron braces with XPDs) is to make a comparison between models with only crumbed rubber added to concrete aggregates in 5,10,15 and 20 % as shown in Table I.

Table I. DEFINITIONS OF 3D MODELS

Crumbed Rubber Ratio in Concrete Aggregates %	0	5 %	10 %	15 %	20 %
<b>Model Code</b>	<b>Type of Braces and Dampers Used in the Model</b>				
3D_Nor_NDR	NA				
3D_Nor_XB_1	1 Middle Chevron Brace without XPDs				
3D_Nor_XB_2	2 Corner Chevron Braces without XPDs				
3D_Nor_XD_1	1 Middle Chevron Brace with XPDs				
3D_Nor_XD_2	2 Corner Chevron Braces with XPDs				

In push over analysis, effect of damping is related to the structural elements material's damping properties and the added damper elements' properties used in the model. Thus, in nonlinear pushover analysis,  $\beta_0$  the hysteretic damping represented as equivalent viscous damping is changing according to materials of structure, damper elements added braces, chevron and XPDs. Also, it is related to  $E_D$  the energy dissipated by damping and  $E_{S_0}$  which is the maximum strain energy.  $\beta_{eff}$  obtained values from analyzing models will be used in estimating spectral reduction factors to decrease the elastic (5% damped) response spectrum. Which will play a main role in the comparison as we will see.

Five types of concrete material were defined in ETABS for the models as crumbed rubber added aggregates according to mentioned percentages hereabove include their damping properties as shown in Table II. Where the values of table were taken by the research done by Najib N. Gergesa, Camille A. Issab, Samer A. Fawazb [5].

Table II. PARAMETERS' VALUES OF RUBBER CONCRETE

Parameters	Normal	5%	10%	15%	20%
Compressive Strength $Mpa$	27.58	17	16	14	12.5
	16	13.82	11.88	8.97	
	26.88	24.13	19.53	13.65	
	30.07	28.15	22.13	16.3	
	39.95	34.63	23.96	18.93	
<b>Average</b>	<b>27.58</b>	<b>25.9</b>	<b>22.78</b>	<b>17.83</b>	<b>14.05</b>
Material Damping Ratio %	18	30	33	38	
<b>Average</b>	<b>5</b>	<b>13.5</b>	<b>21</b>	<b>23.5</b>	<b>28</b>
Young Modulus $Mpa$	24,855	27,000	20,000	17,500	12,500
		30,000	22,500	19,000	14,000
	<b>Average</b>	<b>24,855</b>	<b>28,500</b>	<b>21,250</b>	<b>18,250</b>
Tensile Strength $Mpa$	2.492	1.799	1.789	1.319	
	2.675	2.49	2.295	2.066	
	3.393	3.215	2.837	2.39	

	4.327	3.656	3.015	2.361
	3.22175	2.79	2.484	2.034
<b>Average</b>	<b>2.492</b>	<b>1.799</b>	<b>1.789</b>	<b>1.319</b>

Meanwhile, during analysis the performance curve is being created accordingly, as the pushover analysis depends firmly upon the first mode participation ratio of the structure's modal. The Wen XPDs used in this research were selected due simplicity of manufacturing and availability in local markets. As we saw in Chapter II, XPDs work in both linear and nonlinear behaviors according to the degree of freedom direction in which the link or the damper acts, thus plasticity property of the added links or dampers will add extra energy absorption as a nonlinearity action. To see how all those factors affecting the pushover analysis in structures with multiple damping systems this research was made.

#### A. 3D models specifications used in the research

3D models were developed where the analysis will be done by ETABS Program and are compatible with the following:

- The proposed model will consist of symmetrical 26 stories with 3 bays in X and 3 bays Y directions.
- Wen Plasticity dampers (XPDs) will be add to the 3D models in the middle bays in X and Y directions as first step, next step added (XPDs) will be removed from the middle bays and will be added to corner bays.
- Chevron braces without XPDs will be considered in the analysis as a comparison like it is shown in Table I.
- The first and second step will be repeated for all types of rubber concrete used in this article which explained in previous chapters.
- In 3D models the Response Spectrum will be amplified for

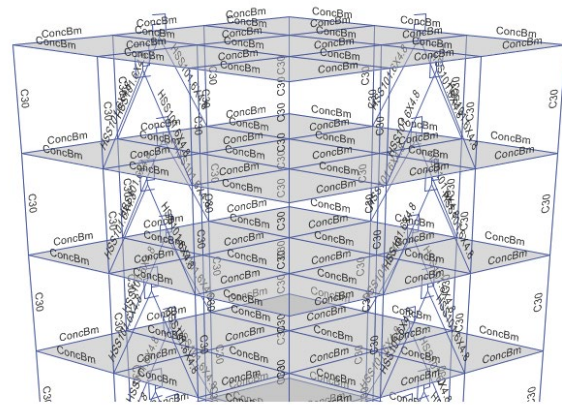


Figure VI. Sections, Bays and Heights of the 3D model with middle XPDs

each model by scale factor to reach the ultimate performance point for each case separately.

Table III. 3D MODEL GEOMETRICAL SPECIFICATION

Parameters	Value	Number
Bay Dimension, X Direction (m)	4	3
Bay Dimension, Y Direction (m)	4	3
Story Height (m)	3	26

Table IV. SECTIONS DEFINITIONS AND LOCATIONS FOR 3D MODELS

Section	Height (cm)	Width (cm)	Location (Story)	Number of elements
C40	40	40	1 to 10	160
C35	35	35	11 to 19	144
C30	30	30	20 to 26	112
Comb	50	30	All Stories	624
HSS101.6*4.8 (Circular)	10.2	4.8	Middle/Corner Bays	208/416
Slab 1 (thin shell)	20		1 to 26	26

Table V. PROPERTIES AND FIELDS OF P-M2-M3 HINGES

Hinge Type P-M2-M3 Moment/ Rotation Data	A	B	C	D	E
Moment/Yield Mom	0	1	1.1	0	0
Rotation/scale factor	0	0	0.01	0.01	0.01

Table VI. ACCEPTANCE CRITERIA OF P-M2-M3 HINGE TYPES

Hinge Type P-M2-M3 Acceptance Criteria	Value (Deformation/SF)
Immediate Occupancy	3.000E-3
Life Safety	9.000E-3
Collapse Prevention	0.01

Table VII. PROPERTIES AND RANGES OF M3 HINGES

Hinge Type M3 Moment/ Rotation Data	Moment/Yield Mom	Rotation/scale factor
E-	-0.2	-0.0457
D-	-0.2	-0.0259
C-	-1.1	-0.0239
B-	-1	0
A	0	0
B	1	0
C	1.1	0.025
D	0.2	0.0269
E	0.2	0.05

Table VIII. ACCEPTANCE CRITERIA OF M3 HINGE TYPES

Hinge Type M3 Acceptance Criteria (Plastic Rotation/SF)	Value (Positive)	Value (Negative)
Immediate Occupancy	0.01	-8.92E-3
Life Safety	0.025	-0.0269
Collapse Prevention	0.05	-0.0457

**Mass Sources:** the mass sources of were models taken from elements self-mass and additional mass and 1 multiplier of assigned super dead loads and 0.25 multiplier of assigned live loads.

While the response spectrum was defined according to American Society for Civil Engineering ASCE7-16 code with those parameters with Tonf, m, C Units.

Table IX. DESIGN RESPONSE SPECTRUM PARAMETERS

Parameters Name	Value
0.2 Sec Spectral Accel, Ss ( $m_{sec}^{-2}$ )	2.29
1 Sec Spectral Accel, S1 ( $m_{sec}^{-2}$ )	0.869
Long Period Transition ( $sec$ )	8
Site Class	C
Site Coefficient, Fa	1
Site Coefficient, Fu	1.3
SDS=2/3*Fa*Ss ( $m_{sec}^{-2}$ )	1.5267
SD1=2/3*Fu*S1 ( $m_{sec}^{-2}$ )	0.7531

## Loads:

- **Dead load** was defined as linear static load and applied for all models as self-weight load for all structural elements in addition to 2 KN/m<sup>2</sup> for each Area element (Slabs).

- **Modal load** was defined as Eigen vectors with zero initial conditions and the defined mass sources of the model.

- **Live load** defined as linear static load and applied as 2 KN/m<sup>2</sup> for all Area elements (Slabs) of the models with zero initial conditions.

- **Push Over Load** which is a nonlinear static typecontinue from the state at end of another nonlinear case which is Gravity nonlinear, where all modal loads applied using modes from the defined modal load with the same mass sources, load application will be Displacement Control with multiple states.

- **Garvity Nonlinear load** is nonlinear static loadcase starting from zero initial conditions (Unstressed State) contains 1 Dead, 0.25 Live and 1 Super Dead load patterns, the load will be applied fully with final state result.

- Load combinations will be taken as the default ETABS 17 combination for concrete RFS (Rigid Frame Structures).

- Analysis Options were taken to be 3D analyzing with advanced solver with auto process.

- Wen plasticity dampers specifications (XPlate Dampers) are functional into U2 direction with nonlinear properties for XPDs which acting in X axis and fixed for all other degrees of freedom, otherwise, all XPDs are functional into U3 direction with nonlinear properties for XPDs which acting in Y axis and fixed for all other degrees of freedom.

Table X. PARAMETERS OF WEN DAMPERS (XPDs)

Parameters Name	Value
n (Number of plates)	8
E (Young Modulus of plates) ( $N_{mm}^{-2}$ )	199947.98
b (Width of plates) (mm)	100
t (thickness of plates) (mm)	8
a (height of plates) (mm)	100
$\sigma_y$ (Yielding strength of plates) ( $N_{mm}^{-2}$ )	248.2113
$\sigma_u$ (Failure strength of plates) ( $N_{mm}^{-2}$ )	399.896
H (rate of strain hardening)	3450.95
d (yield displacement) (mm)	30
Kd (the initial stiffness) ( $N_{mm}^{-1}$ )	6824.89
Fy (yield load) <sub>N</sub>	21180.69
q (yield displacement) (mm)	3.103448
Y <sub>0</sub>	0.413793
P (plastic force for d displacement) (N)	31650.50
Kp (plastic stiffness) ( $N_{mm}^{-1}$ )	1055.016
Kp/Kd	0.154583

## B. Analysis results and comparisons

Figures VII & VIII show that models with corner chevron braces XBs can bear extremely earthquakes with large scale factors of design response spectrum, high ductility and

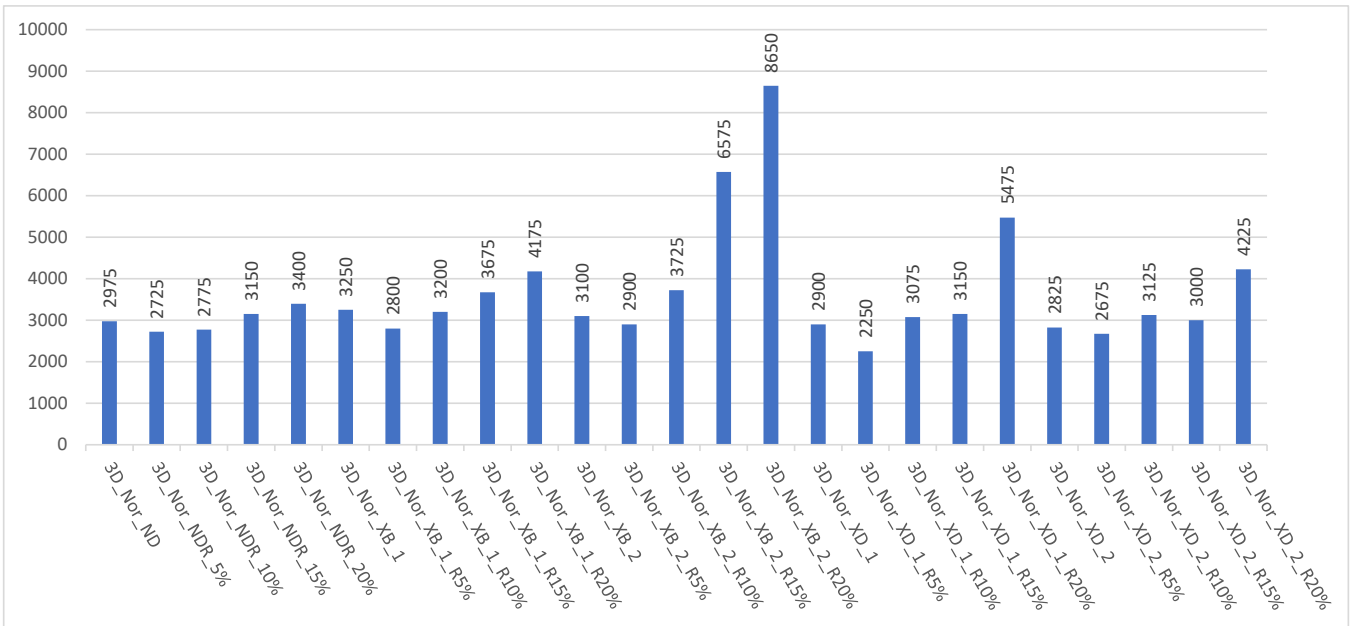


Figure VII. Scale Factor of Applied Response Spectrum Comparison

effective damping ratio for performance point. On the other hand, it shows high value pseudo acceleration and displacement are not preferable in building designs and cause collapsing to nonstructural elements.

response spectrum increases. Comparing models of middle chevron XPDs with corner chevron XPDs, it can be notice slight enhancement in the behavior of middle chevron XPDs models than the corner XPDs ones, whereas, values of effective damping ratio, are slightly higher. In the same time, values of scale factor is slightly lower in the middle XPDs

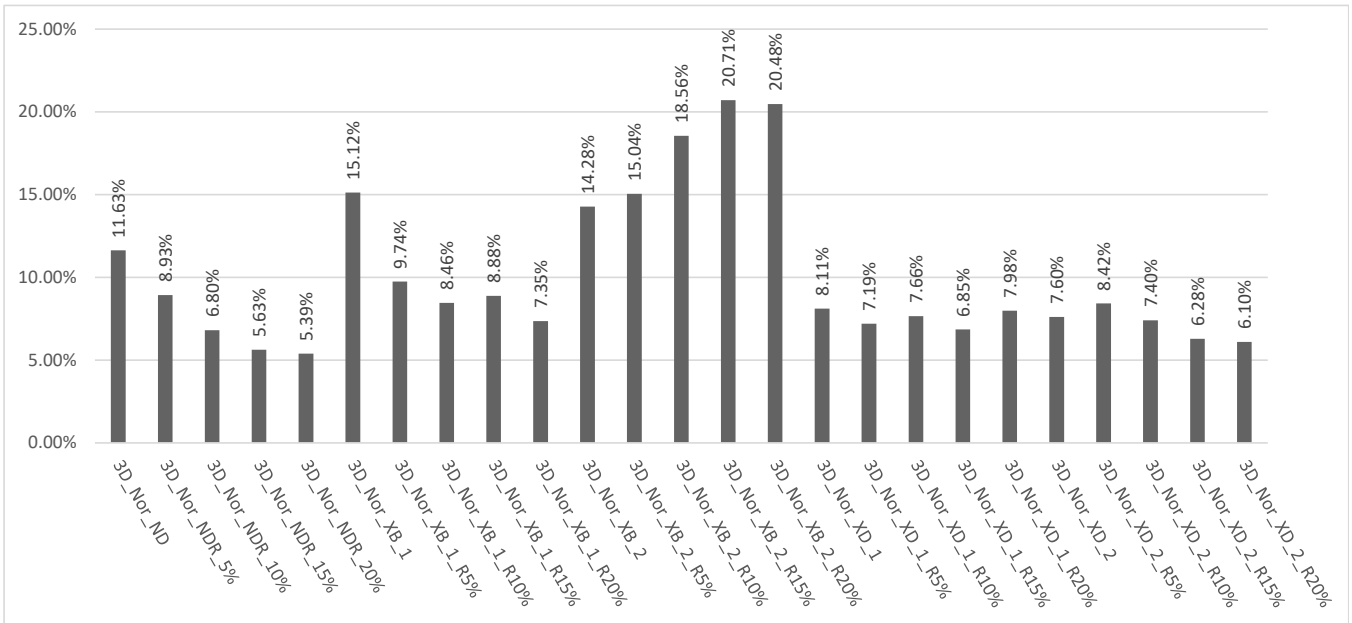


Figure VIII. Effective Damping Ratio Comparison with/out XPDs

For models without braces or XPDs we can see that scale factors of design response spectrum is increasing, meanwhile effective damping ratio is decreasing by increasing crumbed rubber ratio in the aggregates of concrete for the performance point.

Same approach is applied for models that have middle chevron brace without XPDs, yet scale of all comparison parameters is a bit larger than those without XBs. The diagrams show that models with middle and corner chevron XPDs have increasing of effective damping ratio proportionally by increasing crumbed rubber ratio in aggregates, meanwhile scale factor of defined earthquake

models than the corner ones.

As general overview for the results of 3D models, models with corner chevron braces without XPDs -especially for 15% and 20% of rubber ratios- give unrealistic results comparing to other models. Models with middle and corner XPDs can stand stronger earthquakes since scale factor for the response spectrum and effective damping ratios reach maximum values, while transmitted base shear forces, displacement s reach minimum values, especially the models with 10% and

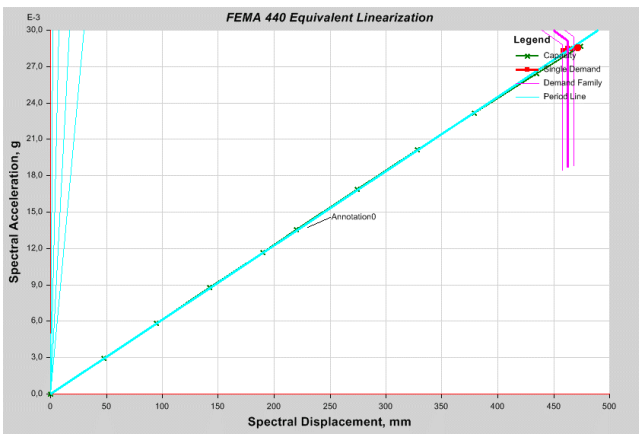


Figure IX. Pushover plot for model 3D\_Nor\_XD\_2\_R15%

15% ratio of added crumbed rubber models. Thus, this leads us to conclude that an optimum situation can be obtained from

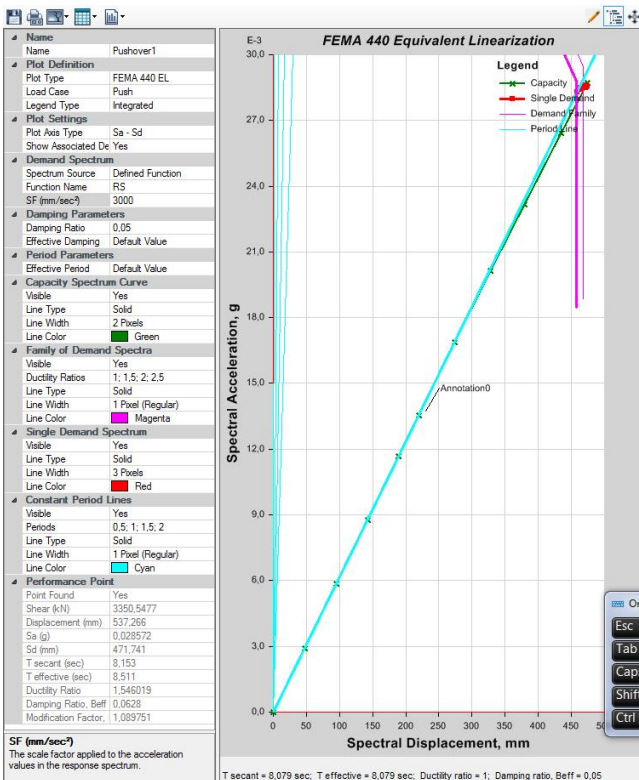


Figure X. Pushover plot with detailed for model 3D\_Nor\_XD\_2\_R15%

mixing the damping systems like adding crumbed rubber to aggregates of concrete, or installing other types of dampers like the XPDs or viscous liquid dampers or mass dampers, etc...

Figures IX and X show that structures in models reach their ultimate capacity with the most reduced demand and highest scale factor of design response spectrum which means that the structures are pushed to the maximum limits.

## V. Conclusion

The XPDs' locations used in the research played a major role in the mechanism of the model's response and behavior towards the design response spectrum curve and also played a major role in raising the scale factor in some study cases compared to other models, in addition to the ratios of crumbed rubber added to aggregates in the concrete used in the models.

Rubber concrete may be a good option to raise the damping ratio of performance point in the structure, it contributes in absorbing energy by a certain amount, but it is not sufficient as individual, adding extra damping systems can form a hybrid system that gives greater effectiveness to raise the effective damping ratio of the structure as a whole, thus greater absorption capacity of seismic energy and increasing scale factor of design response spectra which means increasing structure damping, also proportions of crumbed rubber added should be carefully studied to have the optimum results.

Results obtained are depended on other researches results which may be not accurate, selecting locations of XPDs and ratio of crumbed rubber added to concrete aggregates is critical and depends on geometrical properties of the structural models and the study case. For future researches it is recommended to use optimization applications on such cases to highlight the most proper usage of dampers locations in structures with the best ratio of added crumbed rubber to reach practical applications for such researches.

## References

- [1] Kheir Al-Kodmany (2017). Understanding Tall Building A Theory of Place Making. Routledge Taylor and Francis Group.
- [2] Applied Technology Council (ATC-40 Project) (1996). Seismic Evaluation and retrofit of concrete buildings. Seismic Safety Commission State of California.
- [3] U. D. D. Liyanage, T. N. Perera, H. Maneetes (2018). Seismic Analysis of Low- and High-Rise Building Frames Incorporating Metallic Yielding Dampers, Civil Engineering and Architecture
- [4] SarikaRadhakrishnan, Mr. Sanjay Bhadke. (2016). Seismic Performance of RC Building with X-plate and Accordion Metallic Dampers. International Research Journal of Engineering and Technology (IRJET). Volume: 03 Issue: 07 | July-2016.
- [5] P Sugapriya, R Ramkrishnan, G Keerthana and S Saravanamurugan (2018). Experimental Investigation on Damping Property of Coarse Aggregate Replaced Rubber Concrete, IOP Conf. Series: Materials Science and Engineering 310 (2018) 012003. 10.1088/1757-899X/310/1/01200

# Bidirectional DC-DC Converter Based on Quasi Z-Source Converter with Coupled Inductors

Murat Mustafa Savrun  
Department of Electrical & Electronics  
Engineering, Adana Alparslan Türkeş  
Science and Technology University,  
Adana, Turkey  
msavrun@atu.edu.tr

**Abstract**—This paper presents an improved bidirectional quasi z-source dc-dc converter to achieve a reduced input current ripple. The proposed converter consists of a quasi z-source converter employed coupled inductors, shoot-through switch, and output filter. The proposed converter interface is able to provide bidirectional power flow as well as high voltage gain. Besides, the quasi z-source converter that is equipped with coupled inductors makes it possible to reduce input current ripple. In order to demonstrate the improvement of the input current ripple and verifying power flow functionalities and high voltage capability, a battery connected proof-of-concept model has been developed using MATLAB/Simulink environment. The proposed topology is examined under different power flow directions, battery charging algorithms, and various voltage gain values. Besides, the efficiency of the converter is analyzed for all case studies. The results validate the viability and effectiveness of the proposed converter.

**Keywords**—bidirectional dc-dc converter, quasi z-source converter, high voltage gain, coupled inductors

## I. Introduction

Nowadays, the use of systems equipped with renewable energy sources (RESs), which are rapidly replacing fossil fuels, has been increasing. RESs have a non-linear production behavior due to their nature. Therefore, RESs are often equipped with batteries and DC-DC converters in order to regulate their output. Power electronics converters have great importance in interfacing RESs, batteries and loads in such application areas: renewable energy systems based distributed generation, electric vehicles, microgrids. The step-up converters are frequently used due to the relatively low output voltage levels of RESs and batteries as well as the inherent limitations. Step-up dc-dc converters temporarily store low voltage input energy on magnetic field storage components and transfer it to the output in high voltage levels [1]. The traditional boost converter topology has the advantages of low conduction loss and simplicity of design. However, it is not able to use in voltage-sensitive user applications that need high voltage gain because of the restrictions of limited voltage gain and high output voltage ripple [2]. Several dc-dc converter topologies have been proposed in the literature regarding the high voltage gain issues. The high voltage gain topologies are categorized into two as isolated and non-isolated dc-dc converters [3]. Isolated dc-dc converters are equipped with high-frequency transformers (HFT) in order to isolate the primary and secondary side as well as provide high gain with a turns ratio of it. However, the topologies equipped with high turns ratio HFTs provide limited voltage gain and sacrifice efficiency because of relatively high conduction losses [4]. The non-isolated topologies are frequently used in applications where there is no need for isolation due to their reduced size and cost advantages [5]. In RESs based applications, various high voltage gain dc-dc converters are

applied. These converter types are impedance network [6], switched capacitor [7], capacitor clamped [8], cascaded boost [9], and the quadratic boost [10]. Many of these topologies have some drawback of discontinuous input current [11]. In addition, coupled inductor-based boost converter topologies have been performed to enhance the voltage boost capability with low input current ripple. The converters have high voltage gain thanks to the turns ratio of coupled-inductors. However, higher turn ratios increase the leakage inductance of the coupled-inductors. Therefore, instantaneous voltage spikes and their disruptive effects increase [12]. In [13], a quasi z-source inverter equipped with a coupled inductors topology with the advantages of high voltage gain and reduced input current ripple is proposed to drive a motor.

In this paper, the bidirectional quasi z-source dc-dc converter, which excels with high voltage gain capability, is equipped with coupled-inductors in order to reduce input current ripple. The proposed topology has the superior aspects of each of the two approaches. The high voltage gain is obtained by the quasi z-source converter, while the disadvantage of high leakage inductance due to the high turns ratio of coupled-inductors is eliminated. Besides, the relatively high input current ripple of the quasi z-source converter is reduced via the coupled-inductors. The performance analysis of the converter has been evaluated under different case studies.

The pattern of the paper is organized as follows: The high voltage gain dc-dc converter topology and its operation principle are described in Section II, whilst the control scheme is outlined in Section III. Section IV presents the operation waveforms captured under defined case studies. Finally, conclusions and discussions are presented in Section V.

## II. Power Circuit Structure

The power circuit configuration of the proposed converter is illustrated in Figure I. The proposed impedance network based converter consists of quasi z-source converter, coupled inductors, active switch, and output filter. The quasi z-source converter allows to increase low input voltage to relatively high voltages with its high gain characteristic. The quasi z-source converter is equipped with coupled inductors in order to reduce the input current ripple. The active switch is used to control the shoot-through and nonshoot-through states of the quasi z-source converter. LC filter is used to filter the oscillations at the output of the converter. The quasi z-source converter makes it possible to perform bidirectional power flow. Therefore, the proposed converter has the functionality to be used not only for voltage regulation of RESs with low output voltage, but also for both charging and discharging of low voltage batteries. In order to test and evaluate the power transfer performance in both directions, a simulation model endowed with battery has been conducted.



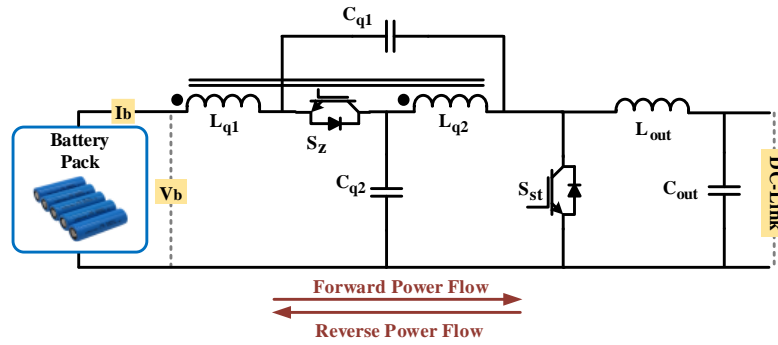


Figure I. Power circuit configuration of the proposed converter

### III. Control Scheme

The controller of the proposed system depends on the duty cycles of two active switches. During the forward power flow, while the freewheeling diode of  $S_z$  switch is biased, the  $S_{st}$  switch is triggered by the determined duty cycle value in order to regulate the shoot-through condition and reach the desired voltage gain value. The duty cycle value varies within the limits of 0-0.5. During reverse power flow, while the  $S_{st}$  switch is in OFF state, the  $S_z$  switch is triggered. The duty cycle value of the  $S_z$  switch determines the charging current of the battery. Batteries are charged according to a charging algorithm in order to extend their service life. The most commonly used battery charging algorithm is constant current (CC) / constant voltage (CV) charging. To perform CC/CV charging, the battery current needs to be able to control. The controller of the converter is illustrated in Figure II.

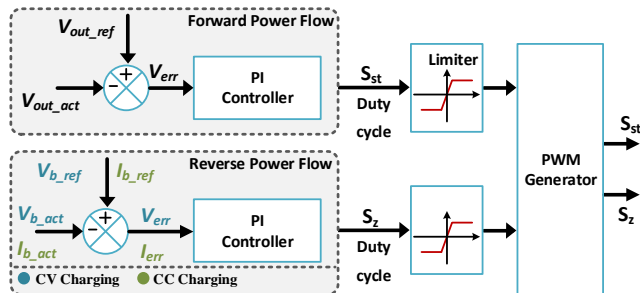


Figure II. Control scheme of the proposed converter

The duty values of the active switches are regulated with proportional-integral (PI) controllers. The controller continuously monitors the output voltage during the power flow towards the load. The desired output voltage level is assumed as a reference value in the controller. The reference and the actual values of the load side are compared and an error signal is produced. The resulting error is given to the PI controller. The controller produces a suitable duty cycle value. During the power flow towards the battery, CC/CV charging algorithm is operated. The controller continuously monitors the battery current. The desired battery charging current value is assumed as reference value in the controller. The reference and the actual values of the current are compared and an error signal is produced. The resulting error is given to the PI controller. The controller produces a suitable duty cycle value. The converter allows to charge battery with its maximum allowable current value (0.5 C) during CC charging. The CC charging switches to CV charging following the battery voltage achieves the determined value. During the CV charging, the charging current gradually decrease to the 0.005

C value. The charging operation ends following the charging current reaches a value close to zero (0.005 C).

### IV. Performance Results of the Proposed Converter

The proposed converter has been modeled, tested and evaluated in Matlab/Simulink environment. To evaluate the performance of the proposed converter, a proof-of-concept model has been developed with a 100 V 50 Ah lithium-ion battery. The parameters of the simulated system are listed in Table I. The switching elements are determined as IGBTs considering the switching frequency and power transfer rating. The performance investigation has been conducted under different voltage gain conditions for forward power flow and under CC/CV charging conditions for reverse power flow. The case studies summarized in Table II has been formed considering all possible operation scenarios.

Table I. Parameters of the simulated system

System	Parameter	Value
Battery	Battery Capacity	50 Ah
	Battery Nominal Voltage	100 V
	Battery Maximum Charge Current	25 A
Quasi Z-Source Converter	Quasi Z-Source Inductors ( $L_{q1}, L_{q2}$ )	1 mH
	Quasi Z-Source Capacitors ( $C_{q1}, C_{q2}$ )	100 $\mu$ F
	Coupled Inductors Turns Ratio	1:1
General	Capacitor ( $C_{out}$ )	200 $\mu$ F
	Inductor ( $L_{out}$ )	1 mH
	Switching Frequency	1 kHz

The first case represents the forward power flow operation under 500 ohm constant load and varying shoot-through duty cycle values. The  $S_{st}$  switch is triggered by the determined duty cycle values in order to regulate voltage gain, while the freewheeling diode of  $S_z$  switch is biased. During the related case, the duty cycle value is gradually increased as: 0.1, 0.15, 0.2, 0.25, 0.3 to evaluate the gain of the proposed converter. To demonstrate the performance of the converter, the input/output voltage/current/ power waveforms are presented in Figure III. Also, the battery discharging current is demonstrated in Figure III to validate the reduced current ripple thanks to the coupled inductors. As illustrated in Figure III, the proposed converter boosts the input voltages to relatively high voltages with different duty cycle values. As evident from Figure III, the efficiency values of the proposed converter are computed as 96.5%, 95.7%, 94.8%, 93.9%, and 92.8% for related time intervals, respectively. The performance waveforms verify the high voltage gain capability with the high efficiency of the converter during forward power flow.

Table II. Operation scenarios

	Case 1					Case 2	
	Time Intervals					Time Intervals	
	0 – 2 s	3 – 6 s	7 – 10 s	11 – 14 s	15 – 18 s	0 – 2 s	3 – 6 s
<b>Duty Cycle</b>	0.1 ( $S_{st}$ )	0.15 ( $S_{st}$ )	0.2 ( $S_{st}$ )	0.25 ( $S_{st}$ )	0.3 ( $S_{st}$ )	0.31 ( $S_z$ )	Decrease
<b>Output Voltage</b>	240 V	407 V	641 V	939 V	1294 V	1000 V	1000 V
<b>Output Current</b>	0.475 A	0.81 A	1.27 A	1.86 A	2.56 A	-2.9 A	Decrease
<b>Battery Voltage</b>	108.6 V	108.5 V	108.3 V	108 V	107.4 V	110 V	110 V
<b>Battery Current</b>	1.1 A	3.17 A	7.93 A	17.22 A	33.27 A	-24.8 A	Decrease
<b>Operation Mode</b>	Gain: 2.21	Gain: 3.75	Gain: 5.91	Gain: 8.69	Gain: 12.05	CC Charging	CV Charging

The second case represents the reverse power flow operation and 0-2 s and 2-4 s time intervals correspond to CC and CV charging operations, respectively. During the related case study, the battery is charged from the load side to verify the reverse power transfer capability of the converter. It is assumed that the dc-link voltage of the output side is 1 kV. While the  $S_{st}$  switch is in OFF state, the  $S_z$  switch is triggered considering the enabled charging algorithm. The battery maximum charging current is determined as 25 A considering the 0.5 times of battery capacity (0.5 C). Thus, the battery is

charged with constant current up to the threshold voltage value. After the battery achieves the determined voltage value, the charging algorithm switches to CV charging. During this stage, the battery current gradually decrease to 0.005 times of battery capacity. Figure IV illustrates the operation waveforms of case 2. The efficiency values for related time intervals are computed as 95.1% and 94.8%, respectively. The performance waveforms reveal that the proposed converter provides bidirectional power flow with high efficiency.

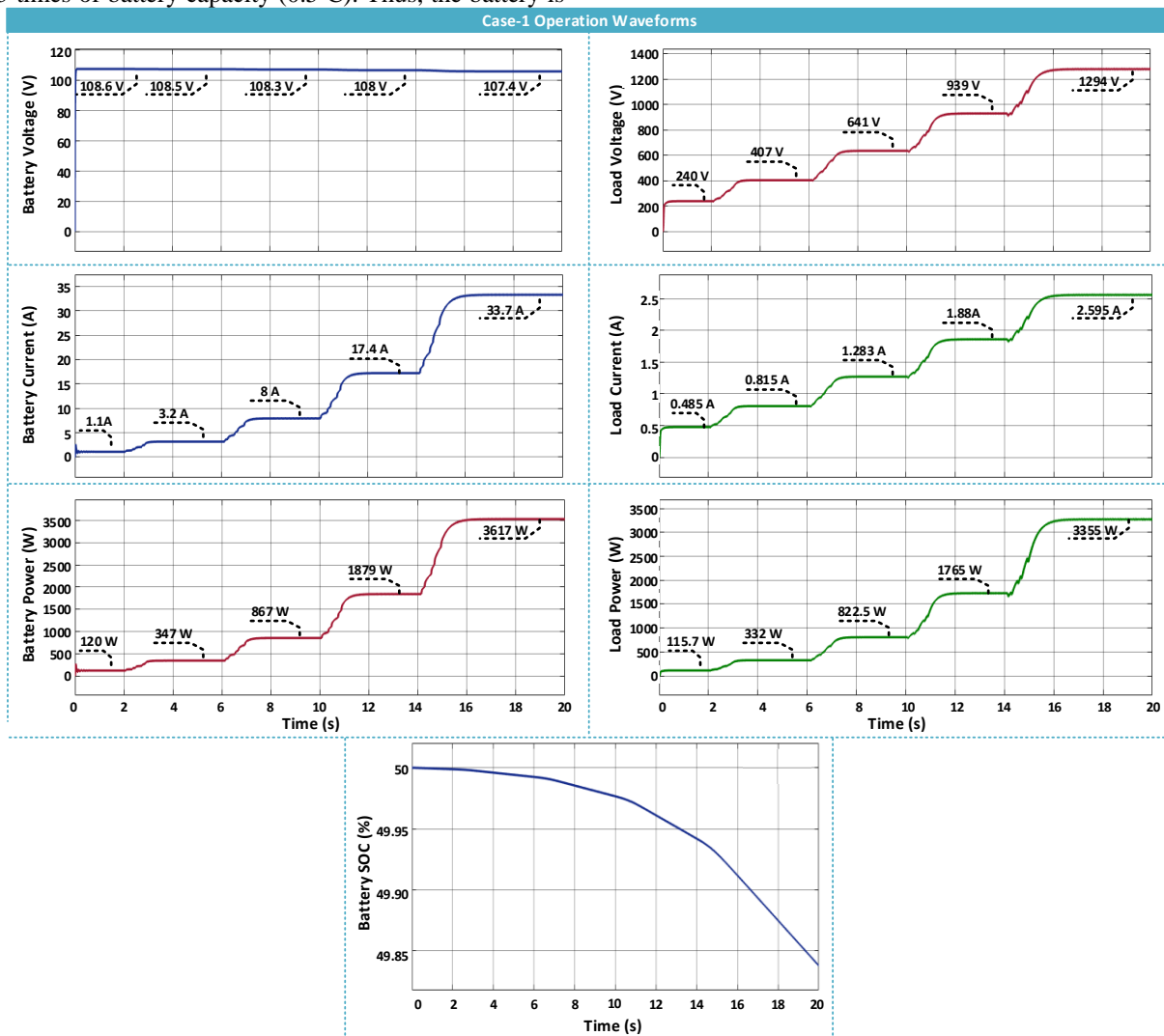


Figure III. Operation waveforms of case 1

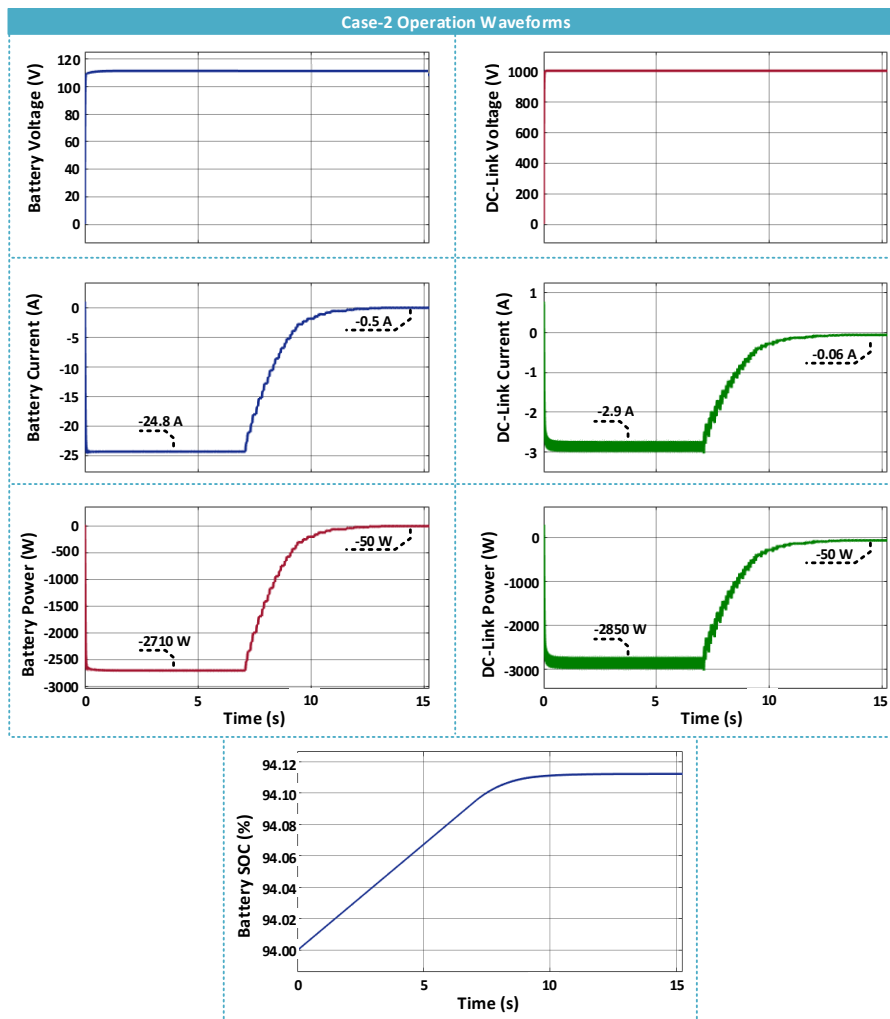


Figure IV. Operation waveforms of case 2

## V. Conclusion

In this paper, an improved non-isolated bidirectional dc-dc converter topology based on coupled inductors equipped quasi z-source converter is proposed. The outstanding features of the proposed converter are; (i) achieving high voltage gain, (ii) providing bidirectional power flow, (iii) mitigating the input current ripple, and enabling precise control and algorithm transitions. Thus, the battery service life is extended with both mitigating the high input current ripple, which is not good for battery life, on the battery side and the performing charging algorithms. The performance of the proposed converter has been evaluated under different operating conditions. The results highlight that the proposed converter is able to provide all functionalities aimed with high-efficiency values.

The proposed converter has a restriction despite its many advantages. The coupled-inductor size increases in parallel with the power rating of the system. Future works will focus on size reduction by increasing the operating frequency of the system.

## References

- [1] M. Forouzesh, Y. P. Siwakoti, S. A. Gorji, F. Blaabjerg, and B. Lehman, "Step-Up DC–DC Converters: A Comprehensive Review of Voltage-Boosting Techniques, Topologies, and Applications," *IEEE Transactions on Power Electronics*, vol. 32, no. 12, pp. 9143–9178, 2017.
- [2] P. García, J. P. Torreglosa, L. M. Fernández, and F. Jurado, "Control strategies for high-power electric vehicles powered by hydrogen fuel cell, battery and supercapacitor," *Expert Systems with Applications*, vol. 40, no. 12, pp. 4791–4804, 2013.
- [3] M. M. Savrun and A. Atay, "Multiport bidirectional DC–DC converter for PV powered electric vehicle equipped with battery and supercapacitor," *IET Power Electronics*, <https://doi.org/10.1049/iet-pe.2020.0759> vol. 13, no. 17, pp. 3931–3939, 2020.
- [4] F. Xue, R. Yu, and A. Q. Huang, "A 98.3% Efficient GaN Isolated Bidirectional DC–DC Converter for DC Microgrid Energy Storage System Applications," *IEEE Transactions on Industrial Electronics*, vol. 64, no. 11, pp. 9094–9103, 2017.
- [5] M. Lakshmi and S. Hemamalini, "Nonisolated High Gain DC–DC Converter for DC Microgrids," *IEEE Transactions on Industrial Electronics*, vol. 65, no. 2, pp. 1205–1212, 2018.
- [6] G. Zhang et al., "An Impedance Network Boost Converter With a High-Voltage Gain," *IEEE Transactions on Power Electronics*, vol. 32, no. 9, pp. 6661–6665, 2017.
- [7] W. Qian, D. Cao, J. G. Cintron-Rivera, M. Gebben, D. Wey, and F. Z. Peng, "A Switched-Capacitor DC–DC Converter With High Voltage Gain and Reduced Component Rating and Count," *IEEE Transactions on Industry Applications*, vol. 48, no. 4, pp. 1397–1406, 2012.
- [8] Y. Zeng, H. Li, W. Wang, B. Zhang, and T. Q. Zheng, "High-Efficient High-Voltage-Gain Capacitor Clamped DC–DC Converters and Their Construction Method," *IEEE Transactions on Industrial Electronics*, vol. 68, no. 5, pp. 3992–4003, 2021.
- [9] P. Upadhyay and R. Kumar, "A high gain cascaded boost converter with reduced voltage stress for PV application," *Solar Energy*, vol. 183, pp. 829–841, 2019.

- [10] S. Lee and H. Do, "Quadratic Boost DC–DC Converter With High Voltage Gain and Reduced Voltage Stresses," *IEEE Transactions on Power Electronics*, vol. 34, no. 3, pp. 2397-2404, 2019.
- [11] M. Moslehi Bajestan and M. A. Shamsinejad, "Novel switched-coupled-inductor quasi-Z-source network with enhanced boost capability," *Journal of Power Electronics*, vol. 20, no. 6, pp. 1343-1351, 2020.
- [12] Y. Hsieh, J. Chen, T. Liang, and L. Yang, "Novel High Step-Up DC–DC Converter With Coupled-Inductor and Switched-Capacitor Techniques," *IEEE Transactions on Industrial Electronics*, vol. 59, no. 2, pp. 998-1007, 2012.
- [13] A. Battiston, E. -H. Miliari, S. Pierfederici and F. Meibody-Tabar, "A Novel Quasi-Z-Source Inverter Topology With Special Coupled Inductors for Input Current Ripples Cancellation," in *IEEE Transactions on Power Electronics*, vol. 31, no. 3, pp. 2409-2416, 2016, doi: 10.1109/TPEL.2015.2429593.

# Influences of urban fabrics onto microclimate assessment within the city of Tirana.

Fabio Naselli  
Department of Architecture  
Epoka University  
Tirana, Albania  
fnaselli@epoka.edu.al

Enkela Krosi  
Department of Architecture  
Epoka University  
Tirana, Albania  
ekrosi@epoka.edu.al

**Abstract**— Tirana, the capital city of Albania, could not escape from that feature that characterizes all the cities of the past socialist regime: sudden and low-governed development process. The study aims to emphasize the problematics and microclimate level differences that coexist to the existing urban fabrics within Tirana, which reduces the comfort of city inhabitants, mainly induced by the post-socialist urban growth. We want to point out about the reduction of the open and green areas and the occupation of free soils by high-rise buildings or informal ones, which have increased the UHI (Urban Heat Island) effect, modifying the local metabolism and by reducing the general urban comfort. The UHI effect is escalated further with the substitution of natural materials by asphalt and concrete. An analytical interpretations of land cover ratios starting by both the land use and land cover analyses has conducted for 2 contiguous urban sites into the city of Tirana, as we selected between diverse typologies of fabrics in the same area of the city. In the meantime, the research shown that an uncontrolled edification may generate different urban environments that do exhibit different microclimatic levels despite their location in proximity one each other.

**Keywords**— microclimatic values, UHI, urban metabolism, land cover, land use, informal urban development

## I. Introduction

Approximately half of the population of the world lives in the urbanized areas and there is the tendency to be increased [1],[2]. This phenomenon has resulted in air, noise and land pollution and consequently changed the microclimate of the urban areas. The change in atmospheric and climatic conditions do affect our mood and activities and even our daily productivity [3]. In the urban areas where temperatures register higher values compared to periphery is commonly known as the Urban Heat Island Effect (UHI). UHI is been affected directly by the change in the wind speed, which results to be higher during the night [4] and increases with the increase of the urbanization and population and recently has resulted to alter by the alteration of the land use and land cover ratios [5], [6]. Greenery as a crucial element that mitigates the UHI effect is rapidly reduced, [7] and has a direct relation to reduction of health problems on humans [8]. The fast expansion of urban morphology with the invention of new and automated construction methods and techniques has resulted in alteration of microclimate values especially the temperature of air for different building blocks. Particularly in Tirana after the decline of communist regime an uncontrolled urban growth spread faster almost to all the city area. This uncontrolled urban sprawl has been spread in such an unpredictable manner that even to adjacent building blocks separated by one main road to be seen a big disparity in terms of land use, land cover, building intensity and so on. The present condition is characterized by substitution of natural surfaces and greenery by impermeable hardscapes such as asphalt and concrete which store a large radiant heat and

possess low albedo values. The high building height of recent constructions has reduced the sky view factor ratios and reduced the ventilation of outdoor areas and wind passages which prevents the cooling of urbanized zones [9] and is reducing the daylight level which causes gloom within building canyons [10]. The aim and objective of this paper is to accentuate the problematics associated and induced by the rapid, uncontrolled urban sprawl which directly influences the microclimate of urbanized zones in city of Tirana. The study intends to raise the awareness of environmental regulation authorities on the related phenomenon and contributing in the improvement of thermal comfort, energy saving issues [11] where the building consumes almost 40% of the energy during the whole lifecycle [12], highlight the need of the green areas and improve the quality of urbanized areas and as an overall will orient the strategies towards sustainable urban development [13].

## II. Influences of urban fabrics

### A. The Study areas

The selected sites to be analyzed in terms of microclimate assessment with regards to their landcover and land use features are located in the city of Tirana, Albania (Figure 1), adjacent to one of the main roads of the city which leads the movement fluxes towards the city center and opposite direction. The buildings' typologies aside the road is of different features and architectural values constructed through a time period of hundred years. The very first blocks that face the road are apartments of 4 and 5 floor compositions accompanied with 2 to 3 story height villas reflecting the influence of the Italian architecture. In the back side of the first facing blocks the architecture and urban morphology has lost its space character by the amateurish and profit oriented interventions which have lowered the values of the outdoor common spaces and reduced even in size.



Figure I. Tirana city

The existing open and green areas have been occupied by new buildings that frequently disobey even the urban planning rules and regulations (Figure II).

City of Tirana is characterized by a Mediterranean climate. It is one of the most wet and sunniest cities in Europe. Tirana is characterized by a diverse urban morphology which is developed by the uncontrolled urban sprawl and rapid extension of the city border.

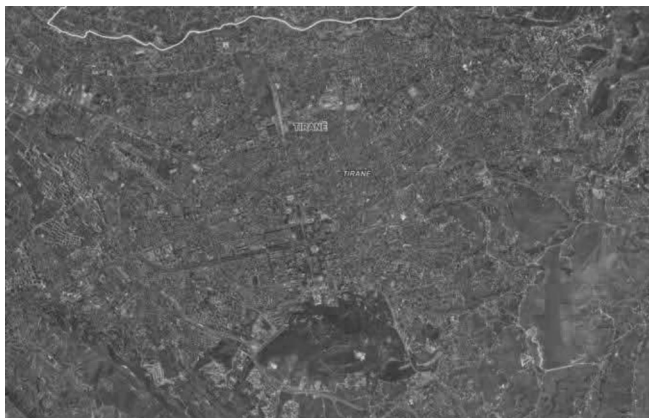


Figure II. Urban growth of Tirana



Figure III. Location of the two study areas (blocks)

The selection of the two blocks along Durrës street (Figure III) and nearby the city center is done intentionally to emphasize the variety of the urban sprawl mainly in the heart of the city for understanding better the whole context and visualizing the problematics associated with the urbanization of Tirana.

### B. The main data used and their interpretation

The data that is collected and has been interpreted in this study is extracted by an analytical work conducted on the land use and land cover of the two selected building blocks. To provide more accurate information from the land covering of the two zones, are used the maps from geospatial portal of republic of Albania, ASIG.

The partial maps that are inclusive for the two selected study zones are furthermore elaborated to extract the intended information for developing the proper comparisons. The land-covering overlapping process with the aerial map is conducted in a pixelated manner or a 5 by 5-meter grid that overlays the aerial map. The 5\*5 m grid is constructed in a manner to simplify the comprehensive and calculation methods, shown also in Figure IV.



Figure IV. Land Covering for Z1 & Z2

One of the axes of the grid is positioned parallel to the main Durrës Street due to the fact that many of the buildings of the zone are positioned parallel to the Durrës Street. This makes the usage of the grid to be more efficient and precise among land cover ratios of the covering materials. Land coverings are grouped or divided into categories such as; buildings, vegetation, and hardscape composed of the asphalt and concrete elements. Grouping is done also within the category of buildings for subdividing them as Buildings with roof tile covering and asphalt or bituminous coverings.

### C. The Land covering ratios

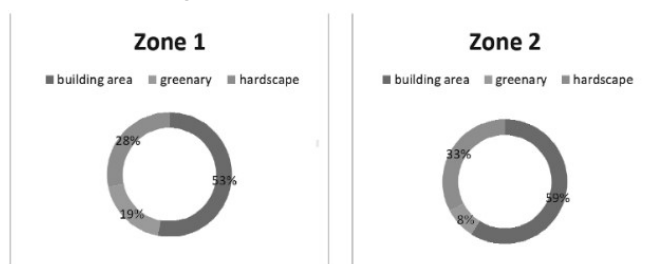
In the Table I. are provided the first data extracted from the calculations on land covering surfaces. There is a clear evidence regarding the differences that exist between the zones despite their adjacent position within the city of Tirana.

Table I. General land Cover Ratios

	Zone 1		Zone 2	
	Area m <sup>2</sup>	%	Area m <sup>2</sup>	%
<b>Total bonded area</b>	18550	100.0%	13650	100.0%
<b>building area</b>	9825	53.0%	8050	59.0%
<b>greenery</b>	3500	18.9%	1125	8.2%
<b>hard scape</b>	5225	28.2%	4475	32.8%

The total bounded zones have a difference in area of approximately 5000 m<sup>2</sup> (Zone 1 > Zone 2). Buildings in the Zone 1 (Z1) occupy 53% of the total land cover area while in the Zone 2 (Z2) occupy 59%. Greenery in the Z1 is covering 18.9% of the total area while in the Z2 is only 8.2%. The hardscape like asphalts and concrete coverings in Z1 stand in the percentile ratio of 28.2 % while in the Z2 32.8%. As an overall comparison between Z1 and Z2 is comprehended that the interventions that induce pollution and increase the UHI effect are much higher at the Z2.

Figure V. Chart of land cover ratios



The ratios of Land Covering can be better understood in the charts.

Table II. Total built area for Z1 & Z2

	Zone 1			Zone 2		
	Area	%	Built tot. area	Area	%	Built tot. area
built area	9825	100 %		8050	100 %	
10k	0	0.0 %	0	675	8.4 %	6750
5k	3575	36.4 %	17875	4525	56.2 %	22625
3k	100	1.0 %	300	1000	12.4 %	3000
2k	3150	32.1 %	6300	1150	14.3 %	2300
1k	3000	30.5 %	3000	700	8.7 %	700
tile cover	4375	44.5 %		1550	19.3 %	
asphalt cover	5450	55.5 %		6500	80.7 %	
			<b>27475</b>			<b>35375</b>

In the Table II. are exhibited the ratios of the buildings by their floor height which are directly proportional to the building intensity. From the Table II. is shown that in Z1 the Land Use coefficient for buildings more than 5 floor are at a value of 36.4% while in the Z2 is 64.6%. The ratio of the inhabited built area over the total Land Use is drastically higher in the Z2 than in Z1, the values are shown in the Table IV. at the intensity values where for the Z1 it is 1.48 and for Z2 2.59. Another issue that is extracted from the Table II. is that materials that store more heat like asphalt and concrete are predominant in Z2. The values shown in Table IV. can be comprehended that by the Coef. Of land usage of Z1=53% and

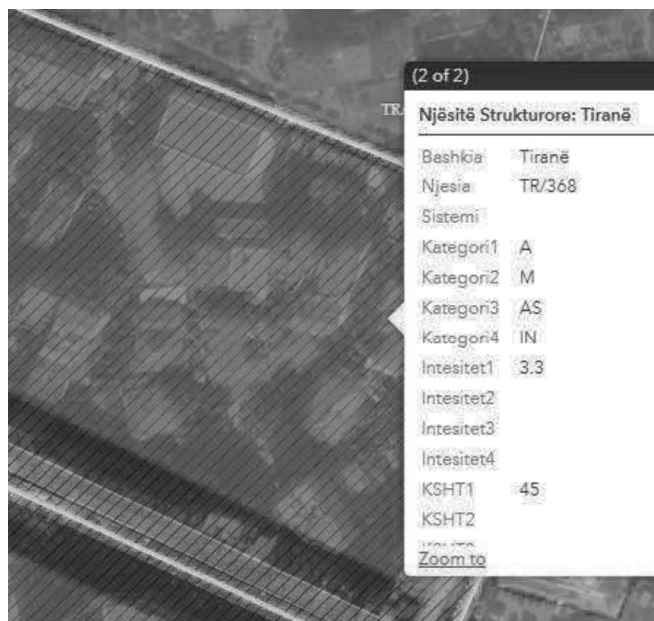


Figure VI. Info of the building regulations for Z1 by municipality extracted from planifikimi.gov.al

Z2=59%, the Z1 does possess higher open and public spaces with green character compared to Z2. With the intensity values is shown that the sky view factor is highly reduced in Z2 due to the large concentration of the buildings composed of more than 5 floors within the zone.



Figure VII. Info of the building regulations for Z2 by municipality extracted from planifikimi.gov.al

Table III. Intensity and L.U.Coeff for Permissibility values and Actual situation

	Zone 1		Zone 2	
	Permissibility value	Actual Situation	Permissibility value	Actual Situation
intensity	3.3	1.48	2.4	2.59
coef. of land usage	45%	53%	45%	59%

In the Figure VI. and Table IV. are shown the values of Land Use and Building Intensity provided by the municipality of Tirana in their building regulation code extracted by the official webpage pplanifikimi.gov.al. In the figure has been understood that Land Use Coefficient is passed with the current conditions for both of the Zones. Regarding the intensity values Z1 has a permissibility of 3.3 while the actual value is 1.48, on the other hand in Z2 the permissibility value of the intensity is 2.4 while the actual condition is 2.59. This has occurred due to the demolition of the Old Italian style villas in the Z2 and construction of the new high rise.

### III. Discussion

Microclimate assessment of urbanized areas is a phenomenon that is widely spread and developed by usage of several different strategies and methods for achieving certain results. Based on past researches conducted on microclimate assessment by usage of various methods and tools results have shown that urban microclimate is directly affected by a number of elements and conditions such as; green ratio, material characteristic, built/unbuilt ratio, sky view factor and so on. Associated to these axioms, our study has been based and developed for understanding the variety of microclimate

change within the urban blocks of city of Tirana. The land use ratios extracted from the two Zones in this research have been selected intentionally to exhibit the disparities that adjacent urbanized zones do possess.

Z1 analysis has shown that the block is surrounded with 5 floor apartment buildings and within the block dominate low rise private houses which offer green space by their private gardens while in Z2 the demolishing of old private houses is substituted by high rise apartments that have reduced the open and green areas within the block.

Table IV. Concluding Table

Zone 1	Zone 2
high green areas	low green areas
low rise building	high rise building
L. U. = 53%	L. U. = 59%
I=1.48	I=2.59
Large areas of tile covering	Large areas of asphalt
high sky view factor	Low sky view factor

Referring to the concluding Table IV. and considering the past researches on similar topics, it has been comprehended that by measurement of Land Use and Land Cover ratios for the two selected zones can be concluded that between Z1 and Z2 should exist a difference in air temperature within the zones, wind flow and ventilation level. Should also exist a difference of UHI values for the both study zones..

#### IV. Conclusion

Tirana as a city which has passed through several urban transformations exhibits different urban characteristics which are evident even in zones located in high proximity within the city. The rapid and uncontrolled urban sprawl has accentuated furthermore the urban morphological variations. This phenomenon has resulted in the alteration of built and unbuilt ratios and reducing the greenery areas which has a direct relation with the increase of urban microclimate temperatures. Reduction of Skyview factor and increasing the amount of heat storing materials in building construction such as asphalt

and concrete has increased the UHI effect which reduces the indoor and outdoor comfort of the citizens of Tirana. The two selected study zones show that even in areas adjacent to each other due to the urban morphological variations and land cover differences exist variations in temperatures and ventilation which has a direct effect in mood and productivity.

#### References

- [1] Wei Ruihan, Dexuan Song, Nyuk Hien Wong, and miguel Martin, "Impact of Urban Morphology Parameters on Microclimate" Elsevier, The Netherlands, 2016.
- [2] Grigoraş Georgiana, and Bogdan Urişescu, "Land Use/Land Cover changes dynamics and their effects on Surface Urban Heat Island in Bucharest, Romania" Int. J. Appl Earth Obs Geoinformation, 2019.
- [3] Olgyay Victor, "Design with Climate" Princeton University Press, New Jersey, 1962.
- [4] Ibidem (1).
- [5] Ibidem (1).
- [6] Ibidem (2).
- [7] Erell Evyatar, David Pearlmutter, and Terry Williamson, "Urban Microclimate. New York 10017" Taylor and Francis, 2011.
- [8] Roberts Hannah, Rosemary McEachan, Tamsin Margary, Mark Conner, and Ian Kellar, "Identifying Effective Behavior Change Techniques in Built Environment Interventions to Increase Use of Green Space: A Systematic Review." In: Environment and Behaviour, SAGE Publications, 2018.
- [9] Tsoka Stella, "Investigating the Relationship Between Urban Spaces Morphology and Local Microclimate: a study for Thessaloniki." Elsevier, 2017.
- [10] Oke Timothy Richard, "Street Canopy and Urban Layer Climate." In: Environment and Behaviour, SAGE Publications, 1988.
- [11] Ibidem (1).
- [12] Kocagil Idil Erdemir, and Gul Koclar Oral, "The Effect of Building Form and Settlement Texture on Energy Efficiency for Hot Dry Climate Zone in Turkey." Elsevier, 2015.
- [13] Mills Gerald, "Progress toward sustainable settlements: a role for urban climatology." Theoretical and Applied Climatology, 84, Springer-Verlag, 2006.



# IoT Based Water Management and Monitoring System for Multi-Resources

Sarosh Ahmad

Department of Electrical Engineering and  
Technology  
Government College University Faisalabad,  
Faisalabad, Pakistan  
[sarosh786a@gmail.com](mailto:sarosh786a@gmail.com)

Sheza Yasin

Department of Electrical Engineering and  
Technology  
Government College University Faisalabad,  
Faisalabad, Pakistan  
[shezay42@gmail.com](mailto:shezay42@gmail.com)

Sajal Naz

Department of Electrical Engineering and  
Technology  
Government College University Faisalabad,  
Faisalabad, Pakistan  
[sajalnaz751@gmail.com](mailto:sajalnaz751@gmail.com)

Amina Batool

Department of BBA Industrial Management  
Government College University Faisalabad,  
Faisalabad, Pakistan  
[aminabatool258@gmail.com](mailto:aminabatool258@gmail.com)

Ali Suqrat Yasin

Department of Electrical Engineering  
The University of Lahore,  
Lahore, Pakistan  
[alисуqrat600@gmail.com](mailto:alисуqrat600@gmail.com)

**Abstract**—This paper aims to manage water distribution in an aligned manner so that everyone will get an equal amount of water without wastage. Without the (programmable logic control) PLC and sensor, the feedback is not obtained in a fast manner. The proposed system is fully automated by connecting it to PLC & SCADA (Supervisory Control & Acquisition), which provides an IoT (Internet of Things) solution. This all proposes a Multi-Resource Control System. Through this project human efforts, time, resources wastage and other complications are reduced. Our idea is to minimize water wastage through a transparent, accountable, and efficient water supply system which in results reduce human efforts, minimize the use of different resources like electricity. So, the goal is to implement an efficient multi-resource water management system in an affordable cost. Thus, we are going to design a setup through which we will investigate various parameters like pH, water level, turbidity of water and manage them by comparing itself to a set benchmark and depending on the population, we will manage the flow meter reading. In this research, we have simulated the salt level testing system, UV (ultra-violet) testing system, and PH testing system and designed a system to manage and distribute water equally without wastage of water.

**Keywords**—Programmable Logic Controller, Supervisory Control & Data Acquisition, Internet of Things, Multi-Resource Control System, water level control, wireless sensors

## I. Introduction

The water on the world's surface is unevenly distributed. Just 3% of the water on the superficial level is useable, the lasting 97% is found in the seas. Of freshwater, 69% can be found in glaciers, 30% underground, and less than 1% is in lakes, waterways, and marshlands [1]. This all concludes that just a single percentage of the world's surface water is usable by the population on earth., and 99% of the usable amount is residing underground. Consequently, water administration and dispersion must be done expertly. Due to the rapid growth of the population, water requirement is increasing day by day which boosts several issues such as water scarcity and shortage. These problems have been increasing swiftly, affecting the home users, agricultural lands, and industrial sectors. The traditional approach of pumping underground water through pumps, tube-well, Petter Engine, etc. contains no proper water management. Therefore, the unnecessary use and waste of water reduce the level of underground water gradually which are causing severe problems to the environment as well as to individuals. The future need is to

utilize water resources efficiently with the water management system. A new method of IoT based with PLC and SCADA, such a framework for water is required which deals with utilization of water. Control engineering has passed through many innovative changes over the last few years. Previously, human beings were the only source for manipulating and commanding any framework [1]. Having troubleshoot helps in analyzing and rectifying an error. Due to reliability, can be used for years without any malfunction [2-4]. Automation is the theme of our proposed project as it plays a vital role in controlling human errors. This system works on different parameters of water like level and flow rate. Using these parameters water wastage and water theft can be avoided. On the grounds, evolvment in technology enhances distinctive methods, observing the economical points in perception [5-7]. Internet of Things (IoT) is a system composed of several branches of mechanical, electrical, computing devices, and wireless technologies that can be employed to achieve water management system requirements. The water pump and pump station can be regularized through Multi-Resource Control System (MRCS), which is an innovative technique. The pumps controller, water level in water storages, and alarming framework are comprised as an integral section for MRCS. Additionally, a 4-state switch will be designed which helps in operating the system manually, automatically, using IoT method, and in Off state. This hierarchy will be driven by an IoT technology, directed by the SMS, Wi-fi, or ringtone which will be accessible from anywhere and at any time [8-10]. The water management system through MRCS using IoT technique may be considered among the modern ways in controlling the wastage of water significantly.

## A. Programmable Logic Controller (PLC)

programmable Regulator is a computerized PC that is utilized to control electromechanical procedures through mechanization. PLC is utilized to control procedures, for example, beguilement rides, apparatus in manufacturing plant, water tank extinguishing in the aviation, filling machine control framework in the food industry, shut circle material shrinkage framework, and different procedures in our day-by-day life. PLC was chiefly intended for multi-inputs and multi-yield forms as shown in Figure I. This further reached out to temperature ranges, invulnerability to

the electrical commotion, and protection from vibration and different effects [11].

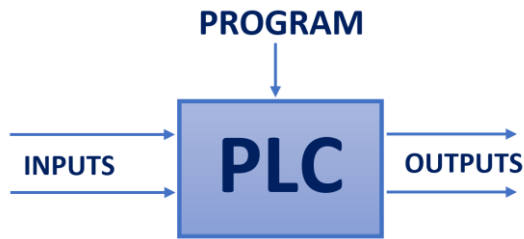


Figure I: Programmable Logic Control System

### B. Supervisory Control & Data Acquisition (SCADA)

SCADA is broadly utilized in the industry for Administrative Control and Information Obtaining of mechanical procedures; SCADA frameworks are currently additionally infiltrating the exploratory material science labs for the controls of subordinate frameworks, for example, cooling, ventilation, power circulation, and so forth. Even more, as of late they were likewise applied for the controls of littler size molecule finders, for example, the L3 moon identifier and the NA48 try, to name only two models at CERN. SCADA frameworks have gained generous ground over the ongoing years as far as usefulness, versatility, execution, and receptiveness with the end goal that they are an option to in house advancement in any event, for exceptionally requesting and complex control frameworks like those of material science tests [17] as can be seen in Figure II.

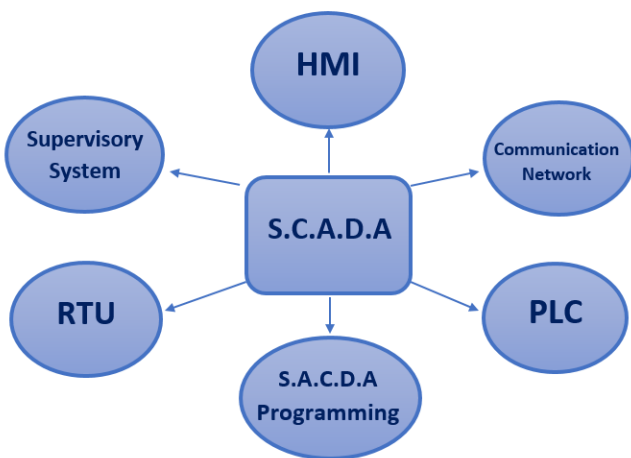


Figure II. Block Diagram of the SCADA system

### C. Internet of Things (IoT)

These interconnected articles have information routinely gathered, broke down, and used to start an activity, giving an abundance of insight to arranging, the executives, and dynamic. IoT characterized as a system of physical items. The web isn't just a system of PCs, yet it has advanced into a system of gadgets of all kinds and sizes, vehicles, PDAs, home machines, toys, cameras, clinical instruments and mechanical frameworks, creatures, individuals, structures, all associated, all conveying and sharing data dependent on specified conventions so as to accomplish brilliant redesigns, situating, following, safe and control and even close to home constant internet checking, online overhaul, process control and organization [19] as presented in Figure III.

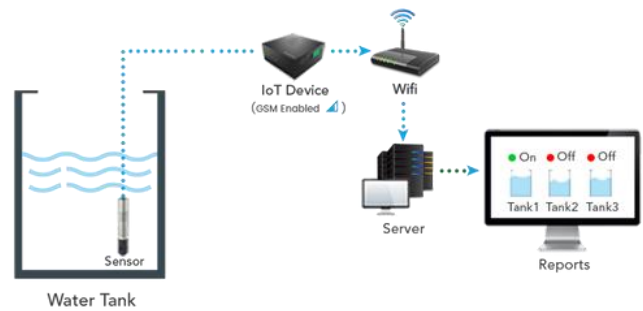


Figure III: IoT Framework in Water Management System

## II. Project Methodology

For making our system efficient, we have added an automation system through PLC. Storage tanks will have some sensors and other instruments, which will be attached to a certain device in order to control and monitor various parameters. These sensors examine physical parameters and convert them into electrical signals in order to give input to PLC. PLC is the main governing party because it controls sensors and other devices and provides data to the controlling room. A controlling room is a SCADA system that stores data in its server coming from PLC and other devices. In current designs, a man must be within the premises of the office for switching on/off the water supply, but we are moving towards PLC which will act like a man for supplying water. The quality parameter which is very vital for healthy water is the pH value. Extreme pH numbers cause severe health issues such as infections to the skin, eyes and also damage different cell membranes. So, in order to avoid all these health and other issues like corrosion of water pipes and mains pH controlling and monitoring system have been installed in our automation system. Proper monitoring of the process is mandatory to have results at an optimal level. SCADA systems have been using in most industries for a long time. It is an efficient system because it shows the information on real-time basis, which helps in sorting the problem and correct them as identified. This SCADA system consists of a primary control center and field sites as per requirement. Point-to-point connections are used for the control center to field site communications. All field sites are interconnected to each other via networking for communication.

This system mainly consists of PLC. This is the central and important part of the system. SCADA system is designed in order to realize the automatic controlling of valve and parameter transformation such as pipeline pressure and water quality [20]. PLC is the heart of our automation system so it provides all logic functions which will be developed through a ladder logic program, used to command PLC. Sensors and Actuators will provide their values and observations to PLC. Then, PLC will monitor and control them based on logic through the ladder program. This logic will be uploaded on PLC through PLC software and can be changed accordingly. PLC is synced with the SCADA system which results in monitoring and commanding the distribution network of water. In the water supply system, we have one storage tank consisting of the level sensor for level monitoring, pH sensor for water quality

monitoring. Other water supply system elements contain pipelines for water flow and pressure switches for closing and opening the outlet valve and the maintenance of volume in the tank. In the case of heavy and high pH water, the amount of chlorine that will be added into the tank is defined in Data acquisition Centre in ppm (parts per million) to make water pure. The equations we used to track ppm in the tank, pressure (PSI), and precipitation (GPM) in the pipe [8]. From the equation, we conclude that the tank's max volume, the Solution's ppm, the elevation height of the reservoir, the pipe diameter from their reservoir, the flow percentage open will be the variables and must be specified at the start.

### A. Automation

Our whole system is based on automation in order to nullify human error and for designing an efficient and modern system. We have used automation for various control elements for operating them automatically for now many years. The foremost gain of an automatic system is that it saves resources, energy, and materials with no compromise and even better results in quality, accuracy, and precision. The block diagram of the PLC system is presented in Figure IV.

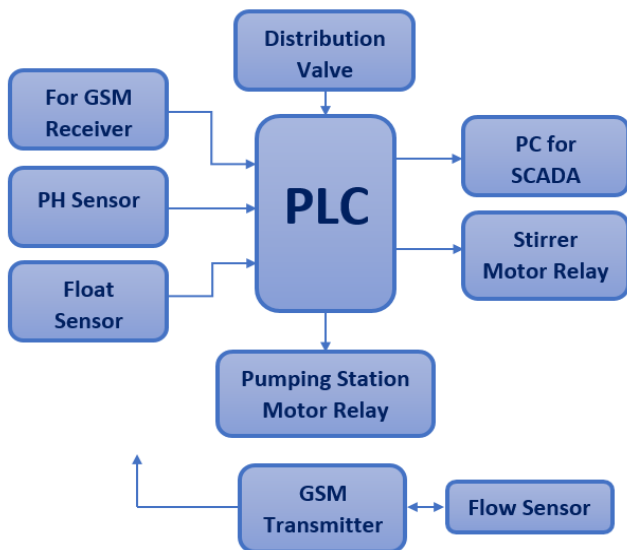


Figure IV: Block Diagram of a PLC System

### B. Pumping Section

In our pumping system design, we have three different sensors for monitoring the water level in the tank. The location of these sensors is specified at the bottom, middle, and top of the tank. They work as a level detector of water as if the sensor detects water level at mid or low, then PLC will get an electrical signal and hence turn on the pumping station. Similarly, the pumping station will be turned off as a water level detector detects a water level at high. For emergency and unforeseen circumstances, we have modeled two motors in the pumping station. From which one will be used in normal routines and the other will be as a backup or for an emergency purpose. We have shown a layout of our pumping section in Figure V. According to our block diagram, we have three processes to perform which are:

- Pumping of a raw or ground and then purifying the input

water.

- Pumping and filtering processes work on six inputs to the input module, such as start push button, stop push button, chlorine tank lower, and higher-level measuring device.
- Allen Bradley PLC controls the process and Wonderware Intouch software SCADA tool is used for monitoring the process.
- After commanding from PLC, the pumping and filtering process will output to the chlorine valve for outlet and reservoirs solenoid valve. In our system design, we have added control for both motors from PLC, which can be commanded as required and the current status and history of the sensors can be seen remotely through the SCADA system.

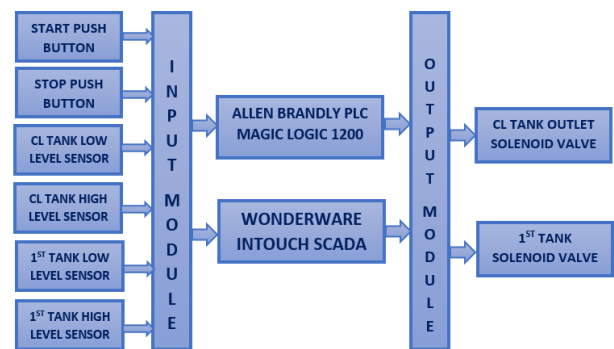


Figure V. Layout Diagram for Pumping Section

### C. Water Treatment/Filtration Section

The second department of the water supply chain is the water treatment/filtration section. Water treatment is very much necessary for avoiding any hazardous circumstances. The water which is focused by us is raw water as it is wasted untreated and furthermore, it has been used in many counties after treatment and filtration processes. So, the model layout of our water treatment system is manipulated in Figure VI. The elements of treating water are dam/underground/raw water, water reservoir tank, flow valve, and pipe. Let assume that the water in the reservoir never runs out. From the equation [8] the elevation of the reservoir with respect to the tank will help to determine the (precipitation) GPM which is flowing into the tank. This result will also assist in analyzing an estimate of (pressure) PSI that is getting out of the reservoir. Hence the GPM calculated will add chlorine to the tank according to the data and its analysis by Data and Control Centre. The same GPM will also be set at the outflow. This will help in analyzing an error and rectifying it in the first instance. This will be done by comparing the GPM at the output with the output if the result matches then there is no error.

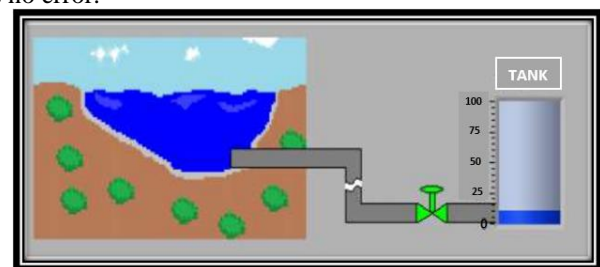


Figure VI. Water Treatment System Tank

### D. Salt Level Testing System

The red color in the first tank is indicating the salt testing. Salt from the water must be removed to get fresh water for humans, irrigation, and for other purposes. Salt is produced as a by-product from the removal of salt from the water of desalination. This removal of salt from water is termed desalination, which produces potential amount by-products from different applications. Figure VII is showing a salt level test simulation diagram. This can be considered as an independent water source. The seas are immensely huge so desalinating them is a very costly process and arises some other big problems for the future. So, the alternatives methods are mostly used.

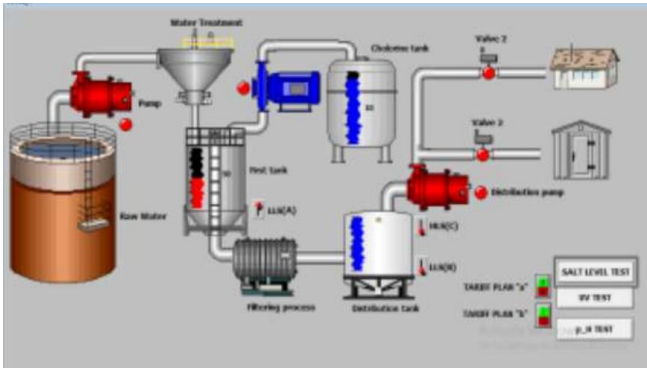


Figure VII. Salt Level Testing System

### E. Ultra-Violet (UV) Testing System

The first tank in green color is indicating the UV testing. For the killing of harmful bacteria and viruses, UV is found more effective. In recent studies by different researchers, it has finally come to know that UV rays are very strong in dealing with bacteria, viruses, and other microorganisms. Test results have been shown in the Figure VIII simulation diagram. The findings show that UV radiation is a good method for the treatment of water for drinking purposes. Some bacteria and viruses have found to survive during tests at high UV doses but are removed from less amount.

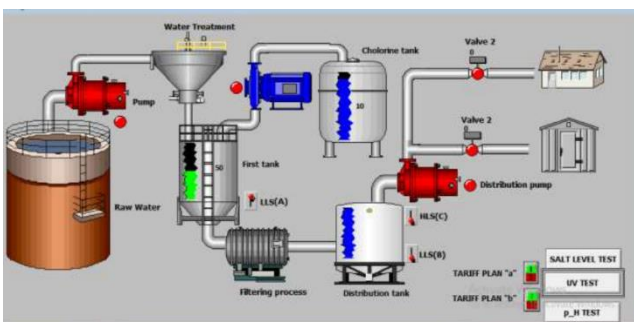


Figure VIII. Ultra-Violet (UV) Testing System

### F. PH Testing

PH is the concentration of ions of hydrogen in water. A pH test is done to analyze the quality of water. As there is no way of maintaining good pH manually so, we have replaced it with automatic control and monitoring for water treatment. In cooling tower systems, pH has been particularly difficult to control manually because the response curve of pH to acid addition is not linear. The first tank in violet color is shown in the simulation in Figure IX. If the color is getting change

it means the process is stopping and then again reprocessing.

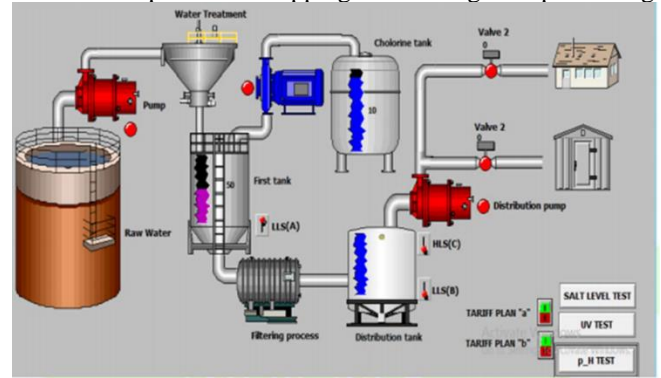
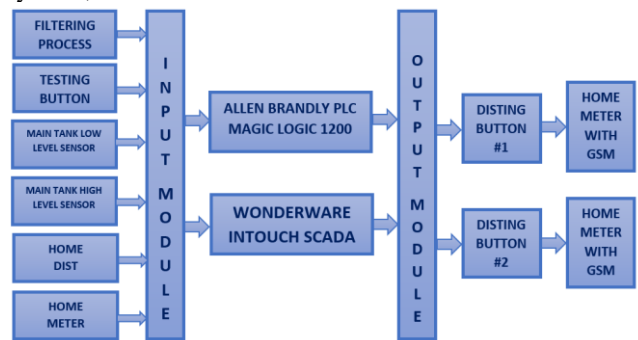


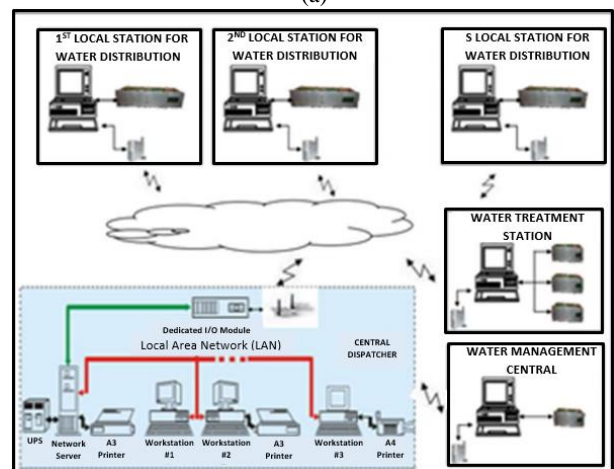
Figure IX. PH Level Testing System

### G. Distribution Section

In our project design, we have considered it as a separate department because we want to have a better water supply for customers and to avoid any faulty conditions. We have seen that most problems have been seen in the water supply/distribution system due to a pressure drop occurrence or pump is used in the home for sucking water directly from the main pipeline passing through the street. We have worked on these main issues and come up with a solution for the water distribution system. In our solution, we will have a control system run by PLC, transmission channel, and other elements with connecting pipes. The block diagram for our proposed design has been shown in Figure X. It shows that all the processes of the water distribution system will be governed by the PLC and for transmitting/receiving control by PLC, a communication channel will be utilized.



(a)



(b)

Figure X. Block Diagram of Distribution Section

In this section, by monitoring the water storing and distribution system, pressure, and other parameters, we can be able to detect any theft happening in streets or somewhere else. Every stage is monitored separately in order to get a precise analysis. If the distribution tank's higher-level sensor is ON, it means that the water distribution is going on. Water can be distributed to all the places at the same time or at different times depending upon requirements and resources. If there is any problem under the distribution control, then the valve will change automatically to manual control and the problem can be rectified. Figures XI and XII is showing the distribution valves simulation figure. The screen is having a status bulb. When the distribution is turned ON, the bulb turns green.

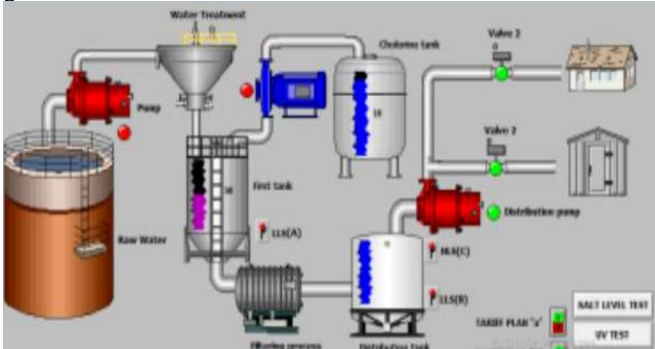


Figure XI. Distribution of Valve 1

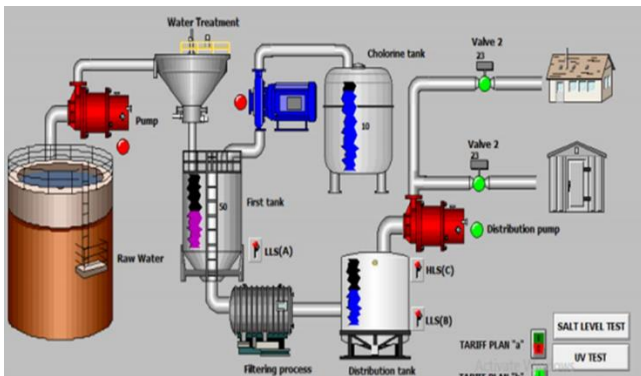


Figure XII. Distribution of Valve 2

### III. ADVANTAGES, APPLICATIONS & LIMITATIONS

#### A. Advantages

Nowadays, PLC has become a basic controller in industries. It is an integral component in industries because it has replaced wiring which provides an efficient way of controlling systems. Other benefits of PLC have been listed below:

##### 1) Flexibility and Reliability

Due to evolutions in technology, one PLC is capable to control multiple systems. Whereas, in the past multiple systems acquired multiple controlling devices. These controllers are very reliable, and the chances of any error are very less as there is very little moving mechanism in the device.

##### 2) Changes and error correction system easier

If any system needs to be modified only changes to program are made. This adds an extra efficiency to the circuit and minimize the time. If using other devices like relays,

alteration in the circuit would be required. This reduces efficiency and increases time.

##### 3) Less Power Consumption

PLC consumes 1/10th of the power as that of an equivalent relay control. Number of contacts in PLC each coil is more than the number of contacts found in relay.

##### 4) Operating speed

PLC operates faster than an equivalent device. Speed of PLC can be determined by units in milliseconds.

##### 5) Reduced Space

PLC is a very compact device because it is a solid-state device rather than electromagnetic and hard-wired devices, where electromechanical motion is required.

##### 6) Ease of Maintenance

Troubleshooting is very much easy because it provides error diagnostic through a program and a software. If the error is found component replacement is also very much at ease.

##### 7) Addition of circuits

It is easy to add multiple circuits to a PLC to provide a control. This can be done without many efforts and saves the amount of money needed to be spent on other controllers.

#### B. Applications

PLC is mostly preferred in certain automation projects. The best market of PLC is industries. Mostly, industries have been using PLC for a long time in their manufacturing processes. These are not only installed for manufacturing purposes but also used in monitoring and automation tasks by industries. PLC requires I/O devices which are then used with different industrial components with respect to their compatibility. In some cases, external circuit design is the requirement for connecting those terminals physically with some programming in ladder logic for making the complete circuit operational. PLC is mostly found in high-end places where the package of PLC is cheaper than the cost of the custom-built controllers. At the low-end level, different automation techniques have been employed to save money, time and effort. The given project's most applications are found in industries where there is excessive use of water with no proper maintenance. This water management system will assist to minimize the useless dissipation of water because it is a vital element on earth. This project also finds its applications in WASA, chemical industries, dying industries and wherever, there is a usage of any liquid for different processes or as a product. This is a one-time investment through which we can save millions of gallons of water annually. The whole system can work remotely through SCADA which can be operated from anywhere and hence will create a safe environment of operation.

#### C. Limitations

- LC is a costly device.
- The project is suitable in industries or in government institutes. Whereas it will not be suitable for local and domestic use.
- Mostly areas are still operating on old controllers so transforming those areas to PLC is a difficult task to achieve.

#### IV. Conclusion

We have worked on water treatment, water level monitoring, water distribution, and controlling system through

automation. Our idea was to establish an economical, feasible, and automated water management system for avoiding the wastage of water. We have worked on this project and proposed multi-resource water management through PLC. In this project, its flexibility is immense as it can be controlled remotely. This is also the subsequent benefit to control water efficiently. The edge PLC has from any other controller or computer is that it is not affected by environmental changes like cold, heat, moisture, dust, etc. Due to the new advancement in technology PLC also upgraded itself by including different controls, motion controls, the capability of networks, divided control, and in some other respects. The remote handling of PLC has been provided through the SCADA system and IoT which is a more modern means. These systems have improved the functionality, performance, and scope of the PLC. Due to the attachment of SCADA with PLC, it gains immense popularity in industries because of remote access.

### V. Future Prospective

Technological trends across the globe are pushing companies and industries towards next an automation-based system, monitoring, and manufacturing era. This all is based on the subsequent work done in previous years to make industries working on automation. Many Inc-operations come forefront and built advanced and simulation tools which are being used in many companies. As our project also has its applications in any automation firm. Moreover, this project has its scope in dyeing industries, petrochemical, pharmaceutical, etc. With the addition of SCADA, the project has gotten huge popularity in remote areas. Remote operations are becoming preferable to simple automation and industries are willing to convert automation tasks remotely which can be accessed from anywhere.

### REFERENCES

- [1] V. C. & L. F. C. Gungor, "Gungor, V. C., & Lambert, F. Research on communication networks of electrical automation system. 50 (7), 877-897." 2006.
- [2] M. D. J. F. & S. M. A. Hadipour, A test set of intelligent control system (MICS) water management system using Internet of Things (IoT). ISA transactions, 96, 309-326., (2020).
- [3] A. A.-S. M. & A. E. A. S. Ali, "PLC water pumping system and frequency control," 2009.
- [4] J. Aziz, "National Water Quality Strategy," Asian Development Bank, 2002.
- [5] S. Rana, "Poor water management is more costly for some countries," Express Tribune, 2019.
- [6] G. & Z. M. H. Murtaza, "Wastewater Production, Treatment and International Use," pages 16-18, May 2012.
- [7] A.-S. & A. Ali, "PLC water pumping system and frequency control," 2009.
- [8] M. I. W. i. t. Silicon Review. (2016, "What is the simple definition of the Internet of Things?" "16 March 2016.
- [9] Hadipour.M, Derakhshandeh.F.J, Shiran.A, "An experimental setup of multi-intelligent control system (MICS) of water management using the Internet of Things (IoT)", ISA Transactions, vol. 96, pp. 309-326,2020,
- [10] Gonçalves R, J. M. Soares J, M. F. Lima R. "An IoT-Based Framework for Smart Water Supply Systems Management", *Future Internet*, vol. 12, no.07,2020.
- [11] Ek, K., Persson, L., "Priorities and Preferences in Water Quality Management - a Case Study of the Alsterån River Basin", *Water Resources Management, Springer*, pp.155-173, 2020.
- [12] Swapnil Namekars, Patel Tayyab Jahngir, Shahid.K. Hannure, Manasi Jagtap, Pratiksha Zagade, "Water Level Controller", *International Journal of Innovative Research in Technology (IJIRT)*, vol. 6, no.11, April,2020.
- [13] M.O. Arowolo, A.A. Adekunle, M.O. Opeyemi "Design and Implementation of a PLC Trainer Workstation", *Advances in Science, Technology and Engineering Systems Journal*, vol. 5, no. 4, pp. 755-761 2020.
- [14] A. Aguilar, M. Pérez, J. L. Camas, H. R. Hernández and C. Ríos, "Efficient Design and Implementation of a Multivariate Takagi-Sugeno Fuzzy Controller on an FPGA," 2014 International Conference on Mechatronics, Electronics and Automotive Engineering, Cuernavaca, 2014, pp. 152-157, doi: 10.1109/ICMEAE.2014.
- [15] S. v. & A. vosough, "PLC and its applications," *International Journal of Multidisciplinary Sciences and Engineering*, vol. 2, 2011.
- [16] "Scheduled Controls Explained (PLC)"
- [17] A. Daneels, W. Salter, "What is SCADA?" *International Conference on Accelerator and Large Experimental Physics Control Systems*, Trieste, Italy, 1999.
- [18] P.M. Adhao, Mahavir's, "Internet of Things (IoT): New Age", *International Journal of Engineering Development and Research (IJEDR)*, vol.05, no.02, 2017.
- [19] K. K. Petal & S. M. Patel, "Internet of Things-IOT: Definition, Characteristics, Architecture, Enabling Technologies, Application & Future Challenges", *International Journal of Engineering Science and Computing*, vol.6, no.05, May 2016.
- [20] G. S. Ashok, "Water Anti-Theft and Quality Monitoring System Through PLC and SCADA," *International Journal of Electrical and Electronics Engineering Research*, pages 355-364, 2013.

# Development of a High Precision Temperature Monitoring System for Industrial Cold Storage

Sarosh Ahmad

Department of Electrical Engineering  
and Technology  
Government College University Faisalabad,  
Faisalabad, Pakistan  
[sarosh786a@gmail.com](mailto:sarosh786a@gmail.com)

Arslan Dawood Butt

Department of Electrical Engineering  
and Technology  
Government College University Faisalabad,  
Faisalabad, Pakistan  
[arslandawood@gcuf.edu.pk](mailto:arslandawood@gcuf.edu.pk)

Usama Umar

Department of Electrical Engineering  
and Technology  
Government College University Faisalabad,  
Faisalabad, Pakistan  
[usama.rwp96@outlook.com](mailto:usama.rwp96@outlook.com)

Sajal Naz

Department of Electrical Engineering  
and Technology  
Government College University Faisalabad,  
Faisalabad, Pakistan  
[sajalnaz751@gmail.com](mailto:sajalnaz751@gmail.com)

Sheza Yasin

Department of Electrical Engineering  
and Technology  
Government College University Faisalabad,  
Faisalabad, Pakistan  
[shezay42@gmail.com](mailto:shezay42@gmail.com)

Amina Batool

Department of BBA Industrial  
Management  
Government College University Faisalabad,  
Faisalabad, Pakistan  
[aminabatool258@gmail.com](mailto:aminabatool258@gmail.com)

**Abstract**—Cold storages are widely used for a number of industries in all over the world, mainly for the food industry. Cold storage facilities play an important role in increasing the shelf life as well as retaining the quality of several raw and processed food items. But there are some problems in old style manually controlled cold storage systems mostly used, which needs to be upgraded with modern technology to reduce potential losses. This research related to the real-time temperature monitoring of cold storage in order to maintain the shelf life, proper monitoring of temperature is required. For the case of apple, if the temperature is maintained between 33.8° to 39.9°F, its shelf life will be from 3-8 months. If the temperature further increases, the shelf life will reduce drastically. In this work, we proposed a highly precise and reliable remote temperature monitoring system to be used in cold storage units. This study constitutes developing an efficient and effective real time remote temperature monitoring system, that will display the accurate and precise temperature on the android app in a cell phone as well as on the LCD. The device was developed using Pt-100 sensor, LT3092, INA-128p, Arduino UNO, Wi-fi Module, LCD Display and PCB. It will provide high efficiency monitoring from remote locations and will greatly help to minimize the temperature variations in cold storages. An economical device with latest features makes this device attractive for industrial use.

**Keywords**—Internet of Things (IoT), surface mount devices (SMD), temperature monitoring, cold storage

## I. Introduction

Cold storage is an important part of the food industries. As stored food items are sensitive to temperature variations and require persistent monitoring. Lack of latest technologies and ignorance about humidity and temperature effects on the fruits results in food safety issues. Slight variations from optimum temperature can cause great economical losses for industries [1]. The main objective of cold storage is to preserve the fruits for a certain period of time. The “Cold storage” is such a storage place where various food and vegetable are stored at cold temperatures for a few months or longer. This allows the food item to be available throughout the year. Every fruit or vegetable has its fix range of temperature for storing them that is known as Optimal temperature range. The temperature of the cold storage should be kept within the Optimal temperature range for proper storage of fruits & vegetables [2-4]. All fruits and

vegetables have their specific shelf life. The “Shelf life” for a fruit or vegetable is the time duration that it can be stored without becoming inadequate for usage, sale, or consumption [5].

## II. Literature Review

The temperature of cold storage must be maintained according to the food stored in it. If temperature of the cold storage is not kept in optimal range, the shelf life of stored items reduces significantly. Temperature and Humidity to be more specific through IoT which doesn't requires the presence of any individual. Temperature variations were monitored precisely and when temperature is above or below the specified range, an alarm is activated [6-7]. At the same time the actuator starts or stops maintaining the temperature in the storage room. The temperature changes seriously affect the quality of farm products. In order to solve the issue, the machinist has to keep an eye on present state of temperature of the cold-storage, even if he is far away. Thus, a remote monitoring system is required by operator to control the temperature automatically [8]. The study designs a remote monitoring system of the temperature. The introduced system is projected to help the operator's facilities and the management of farm products. The method adopted to overcome the problem is use of a diode thermal sensor. The output of the controller is connected to a relay. So, the control method is on-off control. The detected temperature is transferred to the data collection device using serial communication [9]. Various challenges were encountered to optimize control, due to coupling. To decrease the contrary effect of coupling and increase the performance of the refrigeration system of the cold storage, a control strategy with dynamic coupling compensation was considered. On the basis of requirements of the control system, first the dynamic model of the cold storage was established and then the coupling between the components was considered [10]. A fuzzy controller with dynamic coupling compensation was designed to address the challenge. A self-tuning fuzzy controller can serve as the primary controller therefore, an adaptive neural network was adopted to pay for the dynamic coupling. In the end, the control strategy was applied to the refrigeration system of cold storage. The simulations were performed in the condition of a start-up by changing the load and degree of the superheat. The simulation results verified

the efficiency of fuzzy control with dynamic coupling compensation [11].

Our research relates to the real-time temperature monitoring of cold storage. To maintain the shelf life, proper monitoring of temperature is required. The system is designed keeping in mind the optimal storage temperature ranges for standard food items like potato (35°-40°F), for garlic is 30° to 32°F and for apple is 30° to 40°F. As the optimal temperature range for maximum shelf life is 2-3°F in most cases, our system aims to achieve a much higher precision with a much longer temperature probe to display device distance. The main features of research are as under.

- Efficient monitoring of temperature without any fluctuations in the cold room.
- Remote monitoring from anywhere with great precision.

### III. Proposed Methodology

This research was executed in different stages starting from planning to development of hard components and device. Different methods have been used to monitor & control the temperature of cold storage. However, more efficient and remote systems are need of the time, to save industries from great losses & provide fresh and quality products to the end customers. Mostly the optimum storage temperature range for many products is too small that it covers 2-3 degrees Fahrenheit. Temperature difference of 2-to-3-degree Fahrenheit require a very precise measurement. Manual monitoring to this extent is quite difficult for anyone. The Cold storage is a big storage room in which there could be different temperature at different places. At the entrance, the temperature is high compared to the area near the refrigerator. Therefore, the device requires large no. of pt100 sensors at different places to measure the average temperature of the room. The sensor that is placed far apart from the device has a lengthy wire. The wire itself has a resistance that is a prominent cause of error in temperature measurement. Resistance that the device is taking in account in this case, is not the resistance of the sensor only but the resistance of the wire is also added into it. Therefore, the resistance of the wire is a major problem for pt100 sensors. We have to design such a circuit that shall cancel out the resistance of the wire.

#### A. Designed Circuit with INA 128

The resistance of wire is the cause of error in temperature measurements. To remove this error; circuit must be designed by using components such as; pt100, resistances of wire, resistance Ro, two current sources. Two wires are attached with the negative terminal of the Pt100 sensor. One of them contains a resistor. A circuit with two loops that makes Wheatstone bridge is designed. R1 is the resistance of Pt-100. The resistance of pt100 varies with the variation in temperature. It varies from 100 ohm to 101.74 ohm for 32-to-40-degree Fahrenheit. The resistances R2, R3 and R4 are the resistances of wire. Ro is the resistance which is set to 99 ohms. Circuit contains two equal values of current sources. The value of both current sources is set to 10m A. It also contains an instrumentation amplifier for giving gain to the output voltage. The gain for INA128 is set to 85 by setting gain resistance to 595 ohms. The designed circuit diagram with INA 128 is shown in Figure I.

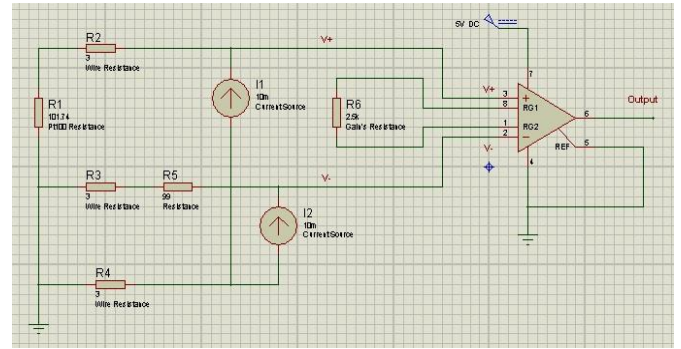


Figure I. Designed circuit connected with INA 128

#### B. Equation of the circuit

By applying the Kirchoff Voltage's Law on the first and the second loop, we have:

Applying KVL on the 1<sup>st</sup> Loop.

$$V_+ = (IR_1 + IR_2 + IR_4) \quad (1)$$

$$V_+ = I(R_1 + R_2 + R_4) \quad (2)$$

Applying KVL on 2<sup>nd</sup> Loop.

$$V_- = (IR_3 + IR_5 + IR_4) \quad (3)$$

$$V_- = I(R_3 + R_5 + R_4) \quad (4)$$

The following equations can be calculated as follows.

$$V = (V_+) - (V_-) \quad (5)$$

$$V = [I(R_1 + R_2 + R_4) - I(R_3 + R_5 + R_4)] \quad (6)$$

As the length of wires for a single sensor is same so we can write  $R=R_2=R_3=R_4$

$$V = [I(R_1 + R + R) - I(R + R_5 + R)] \quad (7)$$

$$V = [I(R_1 - R_5)] \quad (8)$$

This is the final equation of our design. In this equation the resistance Ro ( $R_0=R_5$ ) is subtracted from the resistance obtained from the pt100. This  $V_+$  and  $V_-$  is applied to the Instrumentation Amplifier (INA-128p).  $V_+$  is attached to pin 3 and  $V_-$  is attached to the pin 2 of INA-128p.

#### C. Simulation Results

The circuit is being designed in a Proteus software to get simulated results. Simulation's results at the output pin of the INA- 128 are shown in a Table I having the temperature values with corresponding resistance of Pt-100, voltage difference at input pins of the INA-128 and the voltage at the output pin of the INA-128.



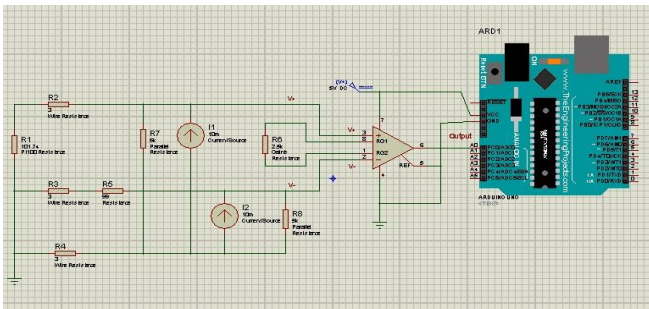
**Table I.** Simulated Results of INA-128

Temperature (F)	Resistance of PT100 (ohm)	Input Voltage to INA128 (mV)	Output Voltage (V)
32	100.00	10	0.85
33	100.22	12.2	1.03
34	100.43	14.3	1.21
35	100.65	16.5	1.40
36	100.89	18.7	1.59
37	101.09	20.9	1.77
38	101.30	23	1.95
39	101.52	25.2	2.14
40	101.74	27.4	2.33

Simulated results show that the voltage at output pin of the INA-128 ranges from 0.85 to 2.33, which must be in range of 0 - 5 volt. It is achieved by setting gain equal to 85.

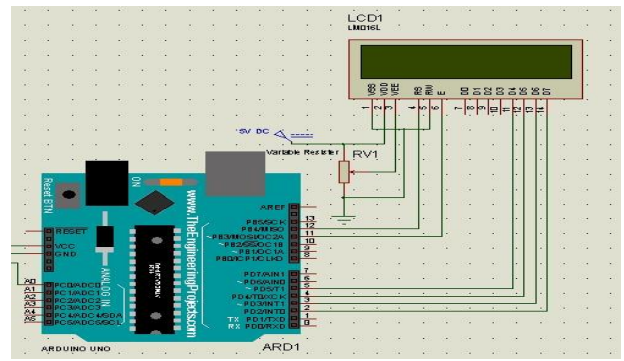
**D. Arduino Interfacing with Proposed Design Circuit**

Arduino is programmed in Arduino IDE software. When the program is installed into the computer by using Arduino IDE then USB cable is used to link the board with the computer. Now opened the Arduino IDE and chose the precise board by selecting Tools>Boards>Arduino/Genuino Uno, and choose the correct Port by selecting Tools>Port. Arduino Uno is programmed using Arduino programming language based on Wiring. Write down the program of the research. To get it started with Arduino Uno board and hardware, load the written program. When the code (also shown below) is loaded into your IDE, we clicked the ‘upload’ button given on the top bar. Once the upload is finished, we observed the Arduino’s built-in LED blinking. This is the design until now with Arduino in it. The output pin of INA has been to the A0. The Arduino is power up by 5V DC source as shown in Figure II.



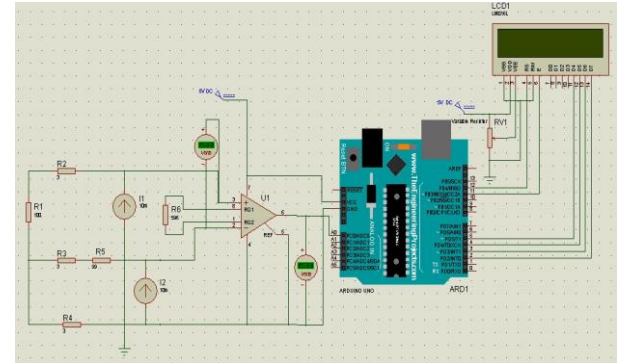
**Figure II.** Designed circuit connected with Arduino Uno

The LCD (16\*2) is connected to the Arduino to get the temperature value at LCD. It is connected in such a way that the 4, 5, 6 and 7 pins of LCD, connect to the 5, 4, 3 and 2 pins of Arduino. VDD of LCD is connected to the 5V DC source. VSS and RW is connected to the Ground. VEE is connected to the variable resistor for any voltage in between 0 to 5 volts. RS is connected to the pin no. 12 and E is connected to 11 pins as illustrated in Figure III.

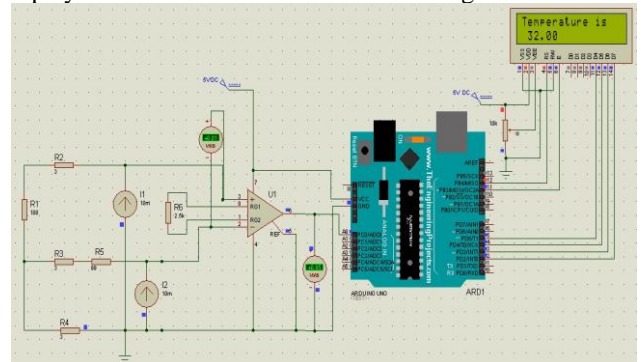


**Figure III.** Arduino connected with LCD

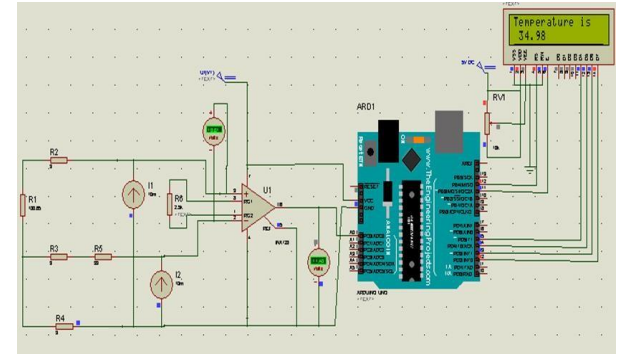
The figure having complete circuit containing Arduino and LCD is as shown in Figure IV.



**Figure IV.** Complete circuit connected with Arduino and LCD  
When the resistance of pt-100 is set to 100-ohm temperature displayed at LCD is 32 as can be seen in Figure V.

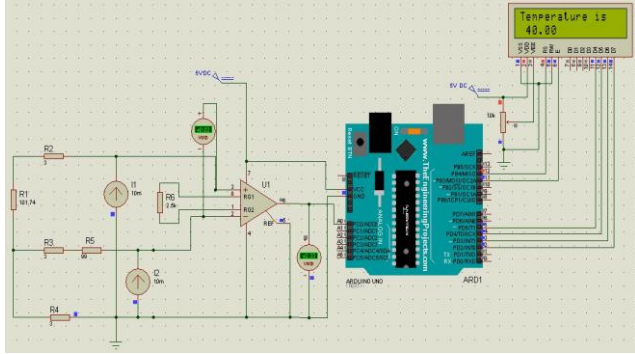


**Figure V.** LCD Temperature Display with resistance of 100 ohm  
When the resistance of pt-100 is set to 100.65-ohm, temperature displayed at LCD is 34.98 as presented in Figure VI.



**Figure VI.** LCD Temperature Display with resistance of 100.65 ohm

When the resistance of pt-100 is set to 101.74-ohm temperature displayed at LCD is 40 as shown in Figure VII.



**Figure VII.** LCD Temperature Display with resistance of 101.74 ohm

#### IV. Hardware Implementation

For the hardware implementation of our design, PCB is used. In our design we used SMD (surface mount devices) which is implemented on PCB. It makes the design more compact. As precise work requires good connections. The appearance will be more professional. We have designed an efficient device having temperature sensing mechanism with a very high accuracy and precision for the cold storages. It will perform remote monitoring by using its android application in the cell phone easily. This device will be robust and will have a battery backup to avoid any interruptions in case of electricity shortage problems. The main purpose of our research is to provide value propositions to the customers, who want to monitor the real time temperature of cold storage remotely.

##### A. Initial Prototype

Initially we have designed a prototype but there were few problems in the PCB of that prototype. First problem was that the pins of INA-128, V+ and Vo, the wire that should be attached to V+ was mistakenly attached with the Vo pin of the INA-128 and vice versa. Secondly in the case of LT-3092 the Vo pins should be soldered with each other as a single pin on PCB but here in the initial PCB both were not soldered but considered as different outputs. The initial PCB prototype is shown in Figure VIII.

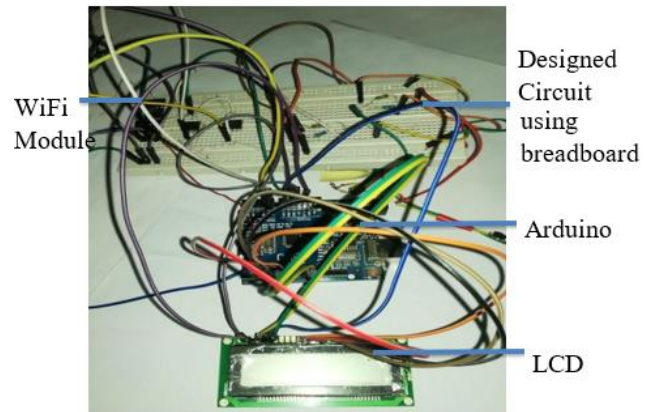


**Figure VIII.** Initial PCB prototype

##### B. Prototype using Breadboard

After the finding the faults in the old PCB we started working on the breadboard because due to the lockdown we were not able to order new PCB. The breadboard circuit is shown in

Figure IX and the final prototype for commercialization is presented in Figure X.



**Figure IX.** Prototype using Breadboard



**Figure X.** Finalized Hardware Prototype

#### V. Conclusion

This research relates to the real-time temperature monitoring of cold storage in order to maintain the shelf life, proper monitoring of temperature is required. We proposed a highly precise and reliable remote temperature monitoring system to be used in cold storage units that will display the accurate and precise temperature on the android app in a cell phone as well as on the LCD. The device was developed using Pt-100 sensor, LT3092, INA-128p, Arduino UNO, Wi-fi Module, LCD Display and PCB. It will provide high efficiency monitoring from remote locations and will greatly help to minimize the temperature variations in cold storages. The main purpose of our research is to provide value propositions to the customers, who want to monitor the real time temperature of cold storage remotely.

#### References

- [1] Karim, A. B., Hasan, M. Z., Akanda, M., & Mallik, A. "Monitoring food storage humidity and temperature data using IoT", *MOJ Food Processing & Technology*, vol. 6, pp. 400-404, 2018.
- [2] Ting, L., & Zeliang, L., "Temperature Control System of Cold Storage", *International Conference on Electromechanical Control Technology and Transportation*, 2015.
- [3] V. C. Chandanashree, U Prasanna Bhat, Prasad Kanade, K M Arjun, J Gagandeep, Rajeshwari M Hegde, "Tinyos based WSN design for monitoring of cold storage warehouses using internet of things", *International conference on Microelectronic Devices Circuits and Systems (ICMDCS)*, pp. 1-6, 2017.
- [4] Ma, X., & Mao, R., "Fuzzy Control of Cold Storage Refrigeration System with Dynamic Coupling Compensation", *Journal of Control Science and Engineering*, pp.1-7, 2018.
- [5] Xu Xiaofeng, Zhang Xuelai, Simulation and experimental investigation of a multi-temperature insulation box with phase change

materials for cold storage, *Journal of Food Engineering*, vol. 292, pp. 110286, 2021,

- [6] Hamid Ikram, Adeel Javed, Mariam Mehmood, Musannif Shah, Majid Ali, Adeel Waqas, "Techno-economic evaluation of a solar PV integrated refrigeration system for a cold storage facility", *Sustainable Energy Technologies and Assessments*, vol. 44, pp.101063, 2021.
- [7] Torres-Sánchez, R.; Martínez-Zafra, M.T.; Castillejo, N.; Guillamón-Frutos, A.; Artés-Hernández, F. "Real-Time Monitoring System for Shelf Life Estimation of Fruit and Vegetables," *Sensors*, vol. 20, pp. 1860, 2020.
- [8] R. Mishra, S.K. Chaulya, G.M. Prasad, S.K. Mandal, G. Banerjee, "Design of a low cost, smart and stand-alone PV cold storage system using a domestic split air conditioner," *Journal of Stored Products Research*, vol. 89, pp.101720, 2020.
- [9] H. Feng, W. Wang, B. Chen and X. Zhang, "Evaluation on Frozen Shellfish Quality by Blockchain Based Multi-Sensors Monitoring and SVM Algorithm During Cold Storage," in *IEEE Access*, vol. 8, pp. 54361-54370, 2020.
- [10] Hina Afreen, Imran Sarwar Bajwa, "An IoT-Based Real-Time Intelligent Monitoring and Notification System of Cold Storage", *Access IEEE*, vol. 9, pp. 38236-38253, 2021.
- [11] Yadav, Ravindra. (2020). Remote Monitoring System for Cold Storage Warehouse using IOT. *International Journal for Research in Applied Science and Engineering Technology*. 8. 2810-2814.

# Modeling and Load Flow Analysis of Electric Vehicle Charging Stations in Power Distribution Systems

Mustafa NURMUHAMMED

Department of Electric and Energy  
Malatya OIZ Vocational School  
Inonu University, Malatya, Turkey  
mustafa.nurmuhammed@inonu.edu.tr

Ozan AKDAĞ

Turkish Electricity Transmission  
Malatya, Turkey  
ozan.akdag@live.com

Teoman KARADAĞ

Department of Electric and  
Electronics Engineering  
Inonu University, Malatya, Turkey  
teoman.karadag@inonu.edu.tr

**Abstract**—As the electric vehicles are becoming part of our lives all over the world, charging them in an efficient way gains more importance as the energy it draws from the distribution network have increased dramatically. Unplanned and uncontrolled charging could cause problems such as transformer overloading, voltage imbalances and power outages in the power distribution network. This paper proposes simulation of charging electric vehicles and comparing power distribution parameters with no load, full load and charging the maximum number of vehicles that system supports. Integrating electric vehicle charging stations to distribution network is analyzed using 11-bus test system. Modeling and load flow analysis are performed and a new test system is proposed. Next, the effect of controlled and uncontrolled charging in the 11-bus test system modeled in this study is discussed.

**Keywords**— *Electric Vehicle Charging Systems, Plug-In Electric Vehicles, Distribution Network, Power Analysis, Modeling and Load Flow Analysis*

## I. Introduction

Electric vehicles are dominating the transportation industry. According to a forecast the percentage of electric vehicles will be more than 25% of entire vehicles in the coming ten years [1]. In general, electric vehicles provide better acceleration, more economy per kilometer driven, less maintenance costs, environmental benefits such as being able to acquire energy from renewable resources and other direct or indirect advantages. These advantages lead to a remarkable rise of electric vehicle awareness and sales around the world. Electric vehicles have expanded at an average yearly rate of 60% in the 2014 – 19 period, totaling 7.2 million cars in 2019 [2]. Everyday use is increasing faster than ever.

Electric vehicles can be charged via relatively slow home AC (Alternating Current) charger, AC charging stations and fast DC (Direct Current) Charging stations. Most home chargers charge at a rate under 3 kW. AC charging stations provide faster charging speeds but onboard chargers of most electric vehicles are designed to charge under 20 kW. DC charging rates fluctuate the most among new vehicles. DC charging rates for new vehicles commonly start at 50 kW and can go up to 250 kW. Overall charging rate of any kind of charging is limited to the lower of the two charging rates; the maximum rate that electric vehicle supports and the maximum rate that charging station supports.

Limited range for electric vehicle owners is becoming less of a concern when planning for long distance journeys as more ultra-fast DC chargers are installed on shorter intervals on strategic locations and highways. Nowadays, 250 kW and 350 kW ultra-fast DC chargers are very common in countries where electric vehicles are adopted in large scale. On the other hand most electric vehicles yet to reach those high speed

charging rates due to design constraints of internal power electronics.

As the number of cars that support high speed charges, and electric vehicles increase, considerable amount of load will be experienced by the power distribution system. This load is considered controlled or uncontrolled load depending on operability. Controlled load is a process of charging or discharging within certain limits or a plan. Uncontrolled load is charging or discharging regardless of any predefined plan, preparation or agreement. The charging of electric vehicles should be planned and rolled out in a controlled manner so that power distribution system components are not overloaded, power quality is not compromised and the system is free of power outages. In addition, the voltage and frequency deviation, harmonics and three-phase voltage unbalance are some of the parameters that affect the quality of the energy in the power distribution network.

This topic is researched in number of studies. One of them is a research utilizing probabilistic analyses in power grid to demonstrate charging effects on the system [3]. Another research observes the effect of charging station loads on reliability indices [4]. In another study, impact of charging electric vehicles on the distribution system is studied using MATLAB/Simulink [5] simulation. In addition, a study group published a report demonstrating specifics of charging electric vehicle effects with probable scenarios, and foresights [6]. One research examined the line status and transformer usage and models were suggested [7].

There are research studies that study reliability of distribution system [8], power quality of high voltage grid [9] [10], voltage deviation [11][12], voltage unbalance [13]. In another study power losses in charging and discharging processes are investigated [14]. One research examine mitigating the instant load increases [15]. Some others are proposing solutions that might reduce the effect of charging by scheduling charge sessions and using charging algorithms [10], [16]. One study proposes smart charging method that optimizes the chargeable power by short term load forecast [17]. In addition, investing in the power distribution system is always a choice; however, the cost associated is high and a research study shows that the top twenty percent of load-serving capacity efficiently utilized only less than five percent of the hours at load duration and serves less than one percent of the electricity demand in the system [18]. Figure I shows the reserve capacity, rarely used peaking capacity and underutilized baseload capacity. Therefore, investing solely to the power network hardware may not be the most efficient solution to protect the power distribution network when electric vehicles are at charge simultaneously.

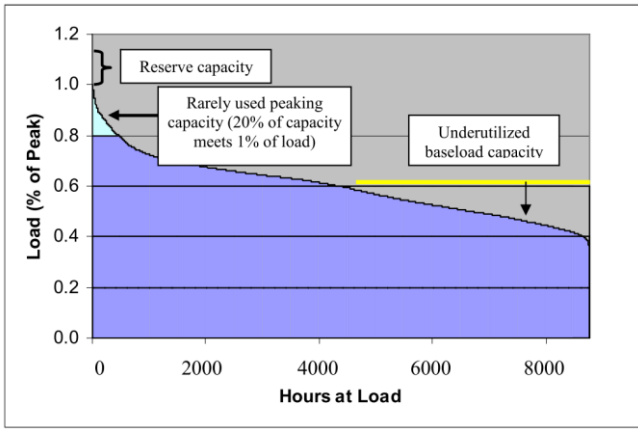


Figure I. Utility load duration graph [18]

Besides mentioned studies, there are many research studies on the impact of electric vehicles on distribution networks [17][19][11][20][21].

buses are 400 V. The line and load data of this distribution system are given in Tables I and II, respectively.

Table I. 11-bus distribution system line data

Bus no	Name	R (ohm)	X (ohm)	B (us)
Bus 2-3	Line 1	0.1051	0.087	373.8
Bus 2-3	Line 2	0.1051	0.087	373.8
Bus 2-4	Line 3	0.052	0.043	186.9
Bus 2-4	Line 4	0.052	0.043	186.9
Bus 3-5	Line 5	0.1051	0.087	373.8
Bus 3-5	Line 6	0.1051	0.087	373.8
Bus 3-6	Line 7	0.063	0.052	224.2
Bus 3-6	Line 8	0.063	0.052	224.2
Bus 3-10	Line 9	0.026	0.021	93.4
Bus 7-10	Line 10	0.1051	0.087	373.8
Bus 7-10	Line 11	0.1051	0.087	373.8
Bus 7-9	Line 12	0.031	0.026	112.1
Bus 7-9	Line 13	0.031	0.026	112.1

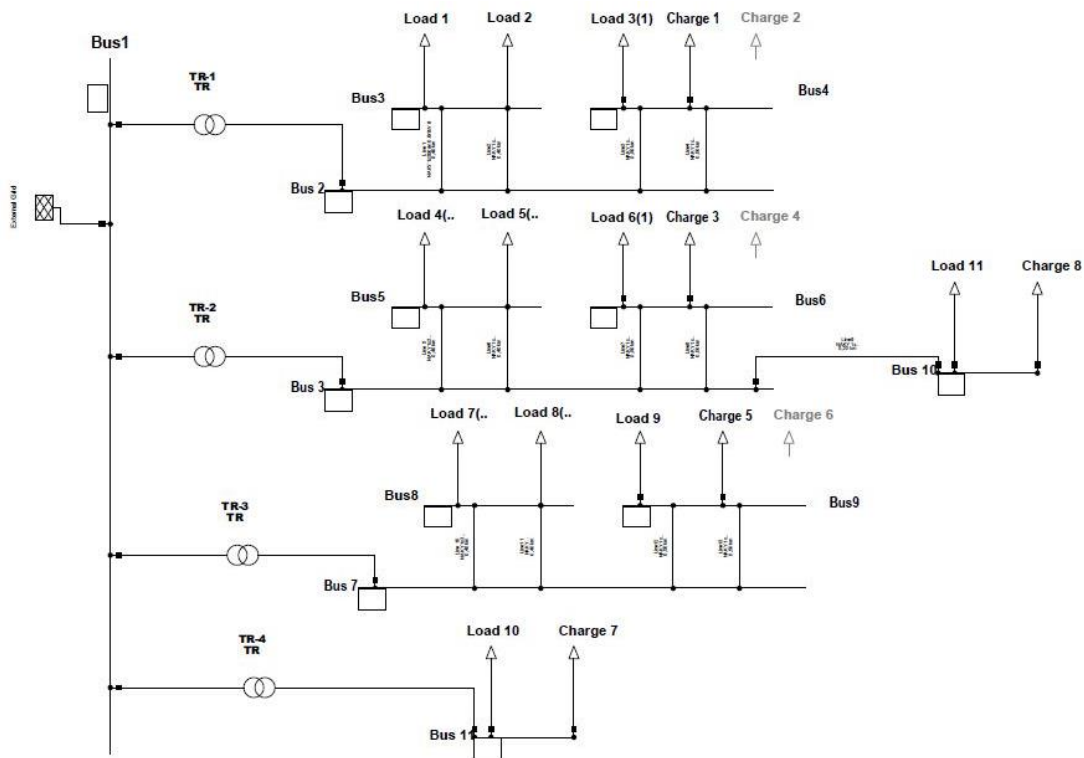


Figure II. Single line diagram of the sample 11-bus distribution system

This paper provides a modeling and load flow analysis to investigate the integration of electric vehicle charging stations to an 11-bus distribution network. Depending on the effects of this integration, additional measures can be taken to mitigate the potential risks on the network.

## II. Simulation

In this study, an electric vehicle is modeled as a 3-phase load. It is assumed to have an average power of 25 kW. An estimated 11-bus cross-section (36 KV/400 V) similar to the power distribution system in Turkey, where electric vehicle charging stations will be connected to, is modeled using DigSilent software [22]. The single line diagram of this model is shown in Figure II.

In Figure II, the bus no. 1 is connected to the mains. Bus no. 1 is the oscillation bus. Bus no 1 is 36 KV, while other

In Figure II, “Load” represents loads connected to the network with the maximum load that can be drawn from the network. On the other hand, the loads specified with “Charge” is the load that electric vehicle charging stations are connected to. In this distribution system, a scenario in which charging stations that support charging 80 electric vehicles simultaneously was formed. All 10 charging stations are designed to be integrated into the charging bus such as Charge1 indicating 10 charging stations. The load model of 10 charging stations is modeled as shown in Figure III.

Figure III shows a load model in the simulation software. Active power, Reactive power, Voltage and Scaling factor parameters are entered for the simulation.

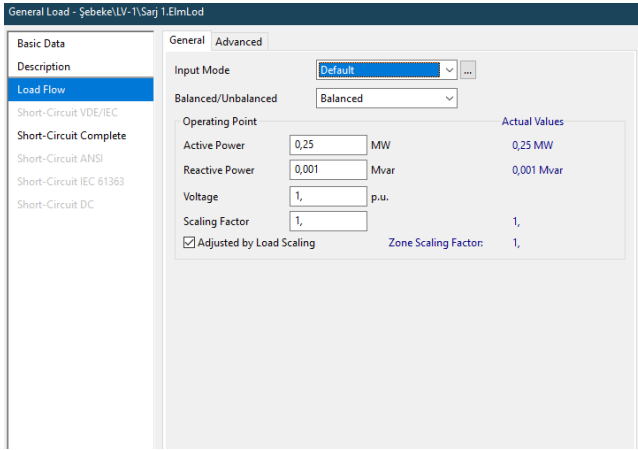


Figure III. Parameter interface of the load model

Table II shows 11-bus distribution system load data.

Table II. Load data of the 11-bus distribution system

Load No	Load		Load No	Load	
	P MW	Q MVAR		P MW	Q MVAR
1	0.245	0.001	7	0.245	0.001
2	0.265	0.0021	8	0.265	0.0021
3	0.245	0.001	9	0.245	0.001
4	0.245	0.001	10	1.02	0.0021
5	0.265	0.0021	11	0.18	0.0021
6	0.245	0.001	-	-	-

### A. Simulation Results

In this power system model, load flow analysis was carried out based on certain cases. In the first case the electric vehicle charging stations are not present. Therefore, by taking these values as a reference, the effects on the grid can be analyzed by comparing whether the charging stations are active or not in use. The load flow analysis can provide information about the loading of lines and transformers and the total losses of the system. The results of the load flow analysis in the absence of electric vehicle charging stations are shown in Table III-IV.

Table III. Total Status

Generation	3.95 MW	0.52 MVAR
Load	3.46 MW	0.02 MVAR
Grid losses	0.48 MW	0.02 MVAR

Table IV. Transformer and line loadings

No	Loading %	No	Loading %
TR1	77.4	Line6	108
TR2	93.1	Line7	41.6
TR3	76.9	Line8	41.6
TR4	85	Line9	59
Line1	108	Line10	108
Line2	108	Line11	108
Line3	41	Line12	40.4
Line4	41	Line13	40.4
Line5	108	-	-

When Table IV is examined, it can be seen that the loading of 4 distribution transformers and lines in the sample power system is normal. Next, taking this information as a reference, a load flow analysis was performed according to the uncontrolled charging case. In this case, all of the charging stations were active. The results are shown in Table V-VI.

Table V. Total Status (Uncontrolled Charging)

Generation	6.32 MW	1.01 MVAR
Load	5.46 MW	0.02 MVAR
Grid losses	0.86 MW	0.98 MVAR

Table VI. Transformer and line loadings (Uncontrolled Charging)

No	Loading %	No	Loading %
TR1	130.3	Line6	109
TR2	174	Line7	150
TR3	124	Line8	151
TR4	106	Line9	151
Line1	108	Line10	109
Line2	108	Line11	109
Line3	143	Line12	131
Line4	143	Line13	131
Line5	109	-	-

When a total of 80 charging stations were commissioned, the first three of the distribution transformers were overloaded. Likewise Line 3-4; 7-8-9; 12-13 are also overloaded. When the other transformer and line loads are examined, it is seen that some of them are at the limit and some are below the overload limit. This may cause damage on the equipment in the power grid and unnecessary power outages. In power systems, lines and transformers can be loaded maximum at 100-110%. Considering the general condition of the power system, controlled charging can be achieved in line with the limits that the power system will withstand. The results of the load flow analysis according to controlled charging case are shown in Table VII-VIII.

Table VII. Total Status (Controlled Charging)

Generation	4.87 MW	0.64 MVAR
Load	4.31 MW	0.02 MVAR
Grid losses	0.55 MW	0.61 MVAR

Table VIII. Transformer and line loadings (Controlled Charging)

No	Loading %	No	Loading %
TR1	100	Line6	108
TR2	100.1	Line7	50.7
TR3	99.7	Line8	50.7
TR4	105.9	Line9	77.2
Line1	102	Line10	108
Line2	108	Line11	108
Line3	88	Line12	84
Line4	88	Line13	84
Line5	108	-	-

With controlled charging, a total of 34 vehicles can be charged simultaneously in this network. Capacity of each charge bus is shown in Table IX. In this case, transformers and lines can be operated within desired loading limits. Thus, in power systems, electrical charging stations should be integrated into power systems within a plan and controlled charging should be exercised. This can be easily done by scheduling electric vehicles in charging stations or distributing power usage among transformers. Alternatively there are other means of keeping the power distribution system safe and stable by applying load flow optimization of energy transmission system using various algorithms such as PSO [23].

Table IX. Number of vehicles that can be safely charged

Charge bus	Number of vehicles that can be safely charged
Charge1	10

Charge3	2
Charge5	5
Charge6	5
Charge7	10
Charge8	2

### III. Conclusion

In this study, the primary goal was to observe the effects of charging electric vehicles to the power distribution system. In order to show the effects of electric vehicle charging stations, 11-bus distribution system was modeled using DigSilent software. The simulation output helps managing integration of electric vehicles, which are spreading rapidly in various countries, into the existing power distribution networks. The differences between when electric vehicles are being charged in a controlled and uncontrolled conditions have been studied. In addition, the adverse impacts of uncontrolled charging are shown and discussed. Furthermore, according to the simulation results obtained from uncontrolled charging, the number of electric vehicles that can be safely charged in a controlled condition can be determined.

This simulation study can help reducing the effects of electric vehicle charging stations by proposing possible impact based on a predefined load scenario. In addition, relatively small 11-bus distribution system is tested and presented with data. Simulation studies allow analyzing the system without taking risk of applying changes on real power distribution systems.

Possible impact of integrating electric vehicle charging stations that simultaneously charge 80 vehicles to an 11-bus distribution network are shown in detail. According to the simulation results, the number of charging stations that can be safely integrated into a distribution system is calculated to be 34. Therefore, a power distribution network can be protected from technical problems and lack of infrastructure by estimating the maximum number of cars to be allowed at the charging stations or limit the energy use that can be drawn from the system.

In the near future, it is expected that more electric powered vehicles especially electric fleet vehicles will simultaneously charge through the power distribution network. Modeling and simulation studies regarding this research topic will play a vital role in planning and designing power distribution grids. Future work can be done in mitigating the possible effects by scheduling electric vehicles in charging stations, distribute power usage or other means of keeping the power distribution system safe and stable.

### References

- [1] G. Giordano, "Electric vehicles," *Manuf. Eng.*, vol. 161, no. 3, pp. 50–58, 2018, [Online]. Available: <https://www2.deloitte.com/uk/en/insights/focus/future-of-mobility/electric-vehicle-trends-2030.html>.
- [2] R. Irlle, "Global BEV & PHEV Sales for 2019," *EV-volumes.com*, 2020. <https://www.ev-volumes.com/country/total-world-plug-in-vehicle-volumes/> (accessed Jul. 07, 2020).
- [3] I. G. Tekdemir, B. Alboyacı, D. Güneş, and M. Sengül, "A probabilistic approach for evaluation of electric vehicles' effects on distribution systems," *2017 4th Int. Conf. Electr. Electron. Eng. ICEEE 2017*, pp. 143–147, 2017, doi: 10.1109/ICEEE2.2017.7935809.
- [4] D. Güneş, İ. G. Tekdemir, M. Ş. Karaarslan, and B. Alboyacı, "Elektrikli araç şarj istasyonu yüklerinin güvenilirlik indisleri üzerine etkilerinin incelenmesi," *J. Fac. Eng. Archit. Gazi Univ.*, 2018, doi: <https://doi.org/10.17341/gazimmfd.416408>.
- [5] B. Yagcitezkin, M. Uzunoglu, and A. Karakas, "Elektrikli Araçların Şarjı ve Dağıtım Sistemi Üzerine Etkileri," pp. 316–320.
- [6] D. Saygın, O. Tör, S. Teimourzadeh, M. Koç, J. Hildermeier, and C. Kolokathis, *Türkiye Ulaştırma Sektörünün Dönüşümü: Elektrikli Araçların Türkiye Dağıtım Şebekesine Etkileri*. 2019.
- [7] M. Kiliçarslan Ouach and E. Çam, "Investigation on the electrical vehicles effects on the electrical power grid," *El-Cezeri J. Sci. Eng.*, vol. 8, no. 1, pp. 21–35, 2021, doi: 10.31202/ecjse.753493.
- [8] H. R. Galiveeti, A. K. Goswami, and N. B. Dev Choudhury, "Impact of plug-in electric vehicles and distributed generation on reliability of distribution systems," *Eng. Sci. Technol. an Int. J.*, vol. 21, no. 1, pp. 50–59, 2018, doi: 10.1016/j.jestch.2018.01.005.
- [9] L. S. Zhao and H. M. Yuan, "The impact of quick charge on power quality of high-voltage grid," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 366, no. 1, 2018, doi: 10.1088/1757-899X/366/1/012033.
- [10] M. Singh, I. Kar, and P. Kumar, "Influence of EV on grid power quality and optimizing the charging schedule to mitigate voltage imbalance and reduce power loss," *Proc. EPE-PEMC 2010 - 14th Int. Power Electron. Motion Control Conf.*, pp. 196–203, 2010, doi: 10.1109/EPEPEMC.2010.5606657.
- [11] K. Clement-Nyns, "Impact of plug-in hybrid electric vehicles on Electricity systems," 2010.
- [12] G. Ma, L. Jiang, Y. Chen, C. Dai, and R. Ju, "Study on the impact of electric vehicle charging load on nodal voltage deviation," *Arch. Electr. Eng.*, vol. 66, no. 3, pp. 495–505, 2017, doi: 10.1515/aee-2017-0037.
- [13] S. Panich and J. G. Singh, "Impact of plug-in electric vehicles on voltage unbalance in distribution systems," *Int. J. Eng. Sci. Technol.*, vol. 7, no. 3, p. 76, 2016, doi: 10.4314/ijest.v7i3.10s.
- [14] E. Apostolaki-Iosifidou, P. Codani, and W. Kempton, "Measurement of power loss during electric vehicle charging and discharging," *Energy*, vol. 127, no. March, pp. 730–742, 2017, doi: 10.1016/j.energy.2017.03.015.
- [15] T. Dragičević, S. Sučić, J. C. Vasquez, and J. M. Guerrero, "Flywheel-based distributed bus signalling strategy for the public fast charging station," *IEEE Trans. Smart Grid*, vol. 5, no. 6, pp. 2825–2835, 2014, doi: 10.1109/TSG.2014.2325963.
- [16] J. De Hoog *et al.*, "Electric vehicle charging and grid constraints: Comparing distributed and centralized approaches," *IEEE Power Energy Soc. Gen. Meet.*, 2013, doi: 10.1109/PESMG.2013.6672222.
- [17] M. R. Poursistani, M. Abedi, N. Hajilu, and G. B. Gharehpetian, "Impacts of plug-in electric vehicles smart charging on distribution networks," *2014 Int. Congr. Technol. Commun. Knowledge, ICTCK 2014*, pp. 1–5, 2015, doi: 10.1109/ICTCK.2014.7033499.
- [18] P. Denholm and W. Short, "An Evaluation of Utility System Impacts and Benefits of Optimally Dispatched Plug-In Hybrid Electric Vehicles," *NREL Rep. noTP-620*, no. October, p. 41, 2006, [Online]. Available: <http://www.nrel.gov/docs/fy07osti/40293.pdf>.
- [19] R. C. Green, L. Wang, and M. Alam, "The impact of plug-in hybrid electric vehicles on distribution networks: A review and outlook," *Renew. Sustain. Energy Rev.*, vol. 15, no. 1, pp. 544–553, 2011, doi: 10.1016/j.rser.2010.08.015.
- [20] K. Clement-Nyns, E. Haesen, and J. Driesen, "Analysis of

the impact of plug-in hybrid electric vehicles on residential distribution grids by using quadratic and dynamic programming,” *World Electr. Veh. J.*, vol. 3, no. 2, pp. 214–224, 2009, doi: 10.3390/wevj3020214.

- [21] J. Balcells and J. García, “Impact of plug-in electric vehicles on the supply grid,” *2010 IEEE Veh. Power Propuls. Conf. VPPC 2010*, pp. 5–9, 2010, doi: 10.1109/VPPC.2010.5729217.
- [22] “PowerFactory.” [Online]. Available: <https://www.digsilent.de/en/downloads.html/>.

- [23] O. Akdag, F. Okumus, A. F. Kocamaz, and C. Yeroglu, “Fractional Order Darwinian PSO with Constraint Threshold for Load Flow Optimization of Energy Transmission System,” *Gazi Univ. J. Sci.*, vol. 31, no. 3, pp. 831–844, 2018.



# Obstacle Avoiding Capabilities for The Drone by Area Segmentation and Artificial Neural Network

Mohammed Majid Abdulrazzaq  
Department of Computer Engineering  
Karabuk University  
moh.abdulrazzaq9@gmail.com

Mustafa Mohammed Alhassow  
Department of Electrical and Computer  
Engineering Altinbas University  
Mustafa.alshakhe@gmail.com

Abdullah Ahmed Al-dulaimi  
Department of Electrical Electronics  
Engineering Karabuk University  
Abdalluhahmed1993@gmail.com

**Abstract**— Obstacle avoidance in unmanned Aerial vehicles is an important task to ensure the mobility and safety of the vehicle. It attracted much attention in recent years, and today we know that most drones are controlled remotely using some wireless technology, such as radiofrequency through remote control, with a telephone. Mobile phone or tablet, making a drone always depend on a user who is giving instructions on what to do and who acts accordingly. The biggest challenge is the development of autonomous air agents that complete missions without having any human intervention. In this paper we propose an area segmentation approach that segments the area into smaller areas and classify those areas into (Safe/Unsafe) which will allow the drone to pass safely through the safe areas and avoid the unsafe or obstacle areas. In comparison with other related works our results shows better performance with respect to time, shape, and path length.

**Keywords**— UAV, obstacle avoidance, Segmentation, Wireless, ANN, Classification

## I. Introduction

Automated robots as well as vehicles have already been used in order to complete missions in risky conditions, for example, tasks in thermal power stations, investigation of Mars as well as in observation of opponent powers in the war zone. Apart from these uses and implementations, there is the advancement of higher intelligence automated aeronautical vehicles that are shortly termed as UAVs for upcoming or future battle in order to decrease manual setbacks. One of the primary difficulties for intelligent unmanned aerial vehicles advancement is basically way of path planning or arranging in adversarial conditions. There have been concerns with path planning that effectively concentrated in the field of robotics. An issue occurs in order for preparation in these applications, planning a path is to discover an impact free-way in a climate where there occur obstacles that are either static or dynamic. Early research focused on both holonomic and non-holonomic systems in kinematic movement issues without considering system dynamics, together with static impediments. Other than numerous outside contrasts, the majority of those strategies depend upon couple of various conventional methodologies that includes roadmap, cell-disintegration as well as potential field.

While relocating deterrents are basically associated with arranging issues, the time dimension is added to the configuration space [1] as well as planning is named as motion-planning or direction/trajectory-planning rather than planning of path. Research has as of late been used in the planning of motion particularly that considers dynamic requirements termed as kino-dynamic arranging or planning [3]. The entirety of previously mentioned path or motion

planning strategies center around impediment or obstacle avoidance issues [2]. Zong et.al [4] proposed an obstacle avoidance scheme for space robots based on the mixed double integer values for predictive control, Zhao et.al [5] proposed a cooperative scheme with transfer learning in flocking swarm of UAVs for obstacle avoidance. Trajectory planning basically termed to the planning of an ideal flight path of the airplane or aircraft between the beginning stage and the closure point, considering components, for example, fuel utilization, mobility, appearance time, flight region, and danger level. Trajectory planning is considered as a significant assurance for the effective finish of a UAV and one of the critical advances for mission planning frameworks. Because of specialized restrictions.

Zhang et.al [6] proposed a trajectory tracking in mobile robots in order to avoid and dynamic obstacles in the robot's path, Padhy et.al [12] proposed a feature extraction scheme from the front view camera of the UAV in order to detect obstacles and avoid them. trajectory planning depended vigorously on manual tasks by experts. With the nonstop turn of events and improvement of the avoidance and control framework and innovation, the precision necessities of a UAV for planning the trajectory are getting increasingly elevated, and artificial path planning has gotten increasingly harder to meet the requirements.

Mendoza-Soto et.al [13] proposed an obstacle prediction scheme in order to predict the obstacles in the UAVs trajectory for efficient obstacle avoidance. With the quick advancement of correspondence innovation, different techniques for identifying flight climate data have arisen endlessly, which makes the data got by the trajectory planners increasingly plentiful. To improve flight precision, the planned trajectory ought to fulfill the requirements of landscape following, territory evasion, danger and threat avoidance while fulfilling the performance limitations of the airplane.

The idea of fully automated systems is the absence of a human factor and thus the prevention of human losses in the event of accidents, or much faster decision-making. Despite the fact that almost fully automated systems already exist and their functionality is very accurate, it will take some time before there will be fully automated and trusted robots or vehicles.

In this paper, we train a deep learning algorithm to classify area segments into safe and unsafe areas also creating a safe path that will lead the drone to move from the start point to its goal. The rest of the paper is organized as follows: Section 2 is where we explain our proposed scheme, Section 3 is where we show the simulation and results for our model, Section 4 is where we conclude our work.

## II. Proposed Scheme and Properties

To get a drone to make decisions, first of all, it must be considered that the first problem occurs when it does not have knowledge of the environment surrounding it, implying that the air vehicle does not process any information in this regard. All the information generated by the sensors it carries is sent to the control station or application for its interpretation. Depending on these information the environment will be classified in order to create the safe path as shown in the figure below:

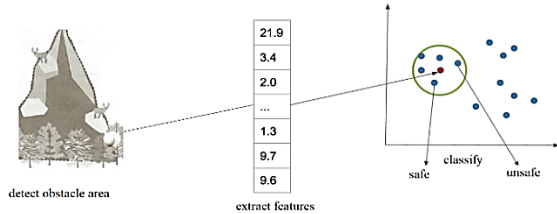


Figure I. Proposed obstacle avoidance scheme.

There is no decision support system independent of the ground control station that would make it adopt a proactive control model. To make this possible, visual perception techniques have to be integrated into the control of UAVs, in order to increase their navigation and direction skills. Yet we do not simulate these perceptions, we merely suggest that in our simulation the sensing device connected to the drone made it possible for the drone to be fully aware of its surroundings.

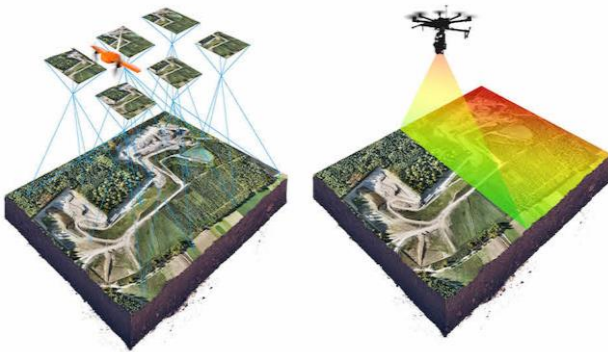


Figure II. General area sensing approach in UAVs[15].

After we acquire all the sensing data from the drone's sensing device, we segment the map of the flying area into groups or smaller areas based on the similarities and distance from one area to another, we use the k-means segmentation algorithm and the Euclidian distance, as shown in the figure below:

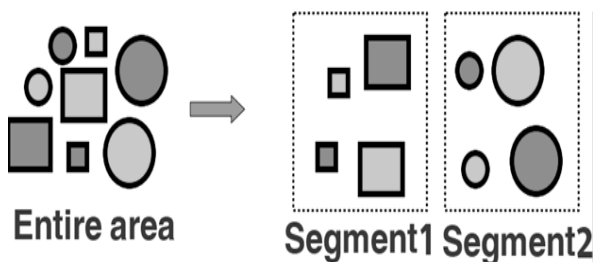


Figure III General area segmentation approach.

## A. K-Means Algorithm

It is one of the most simple algorithm variations (k-means cluster) that is simple to implement and is useful for segmentation of simple image. When it comes to image segmentation, the concept we're talking about is that of image vectorization. It's done by first color quantizing an image and then reducing the available color palette to a certain colors finite amount. During the partitioning of the image, The picture becomes multiple sections. This makes it easier to use various image processing algorithms. The next task of our image segmentation is to locate objects within the original image, as well as to find their boundaries that define the areas of objects that are located. The focus of this particular technique is to locate those pixels with the same hue parameter value. Since these pixels have been grouped into multiple labeled segments, they are commonly referred to as segments.

The Euclidean distance formula has a somewhat more general representation as follows:

$$(L)^2 = \sum_{i=1}^n (x_i - y_i)^2 \quad (1)$$

Where the value for distance of Euclidean is  $L$ ,  $x$  and  $y$  are the two vectors being considered.  $N$  - dimensions number (i.e. vectors components) ( $n = 2$ ). In this instance, only two or three-dimensional vectors are involved, so the following formula is simplified: Let's use this variation of the Euclidian distance formula to calculate When two pixels have a particular coordinate, the distance between them is specific coordinates:

$$L = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2} \quad (2)$$

The  $L$ -value represents the Euclidean distance ( $x_1; x_2$ ) and ( $y_1; y_2$ ) the ( $X; Y$ ) coordinates of two pixels. Then we extract the features of each area in order to classify it into safe and unsafe areas, we use CNN or Convolutional neural networks to extract the features of the smaller areas and put them as an input for the ANN input layer in order to classify them CNNs have assumed a significant job throughout the entire existence of profound learning. They are a key cause of effective utilization of AI.

## B. ANN Applied Features

They were likewise a portion of the principal profound models that functioned admirably, well before discretionary profound models were viewed as suitable. Convolutional systems were additionally a portion of the principal neural systems to understand significant business applications and stay at the bleeding edge of business profound learning applications today. Convolutional systems were a portion of the principal profound working systems prepared with back engineering or backpropagation.

It isn't totally clear why convolutional systems were effective when general converse proliferation systems were considered to have fizzled. It might essentially be that convolutional systems were more computationally proficient than completely associated systems, making it simpler to run various tests with them and tweak their execution and hyperparameters, the figure below shows a general idea of the CNN feature extraction:

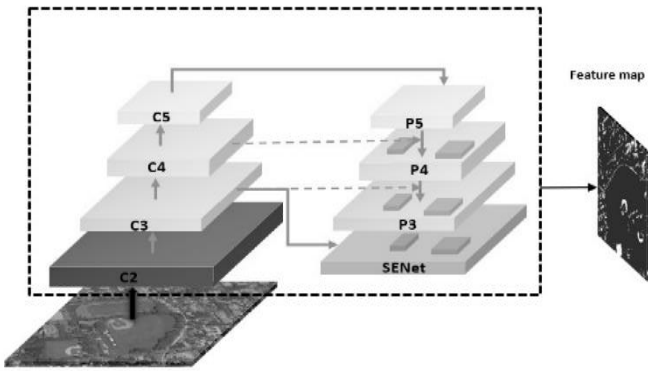


Figure IV. CNN Feature Extraction.

Then we use the CNN features and put them as input values in the ANN classification process, the ANNs perceptions or (input neurons) number of features is same number of each area segment. The features will be processed through an activation function in the hidden layer neurons and output the classification result of that area as shown in the figure below:

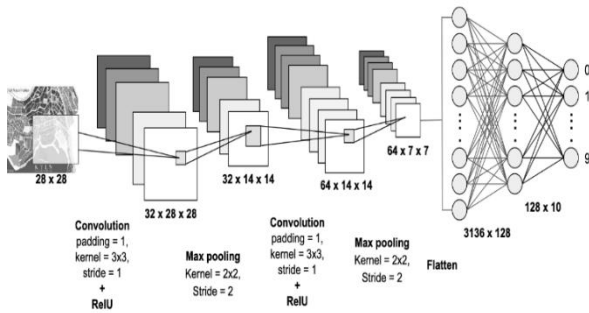


Figure V. Proposed CNN-ANN scheme.

### III. Evaluation and Simulation Results.

To simulate our method in a 3D view, we use the Unity 3D environment combined with the codes for the CNN, ANN, and K-Means compiled in Visual studio, we set up an environment of 3D space with 3D obstacles and set the start and destination point for the UAV as shown in the figure below:

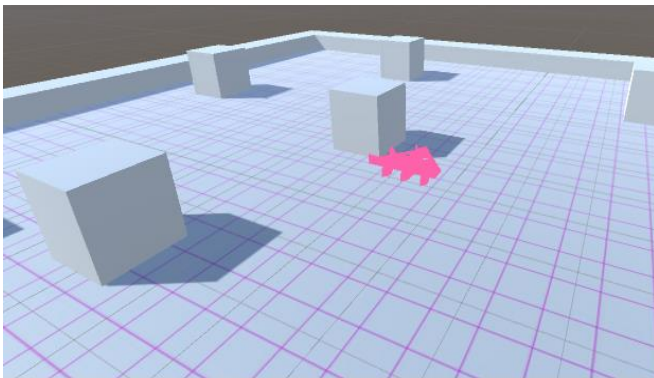


Figure VI. 3D environment set up.

Then we segment the area of 3D space into smaller areas using the K-means algorithm, the K-means will take into consideration the distance and similarities between the segments, the obstacle or the greater number of obstacles in an

area will be the center of the cluster (segment) as shown in the figure below:

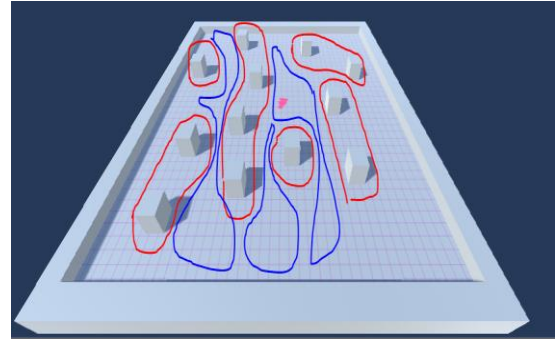


Figure VII. Area segmentation by K-means algorithm.

Then CNN will extract the features of each segment and export them to ANN which will take those features and classify the areas into Safe and Unsafe segments in order to set a path between the safe segments and avoid the unsafe segments. To evaluate our method, we use the mean square error index as a cost function for our classification, normally the ANN has a high accuracy rate and low MSE, the equation for the MSE is shown below:

$$MSE = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n} \quad (3)$$

The objective of the cost functions is to indicate if the output of the predictions corresponds to the expected values and by adjusting the weights by the optimizer the error at the output of the network is reduced. In order to measure the error of the output of the predictions, that is, to measure whether the network was learning or not, it was made use of the function of mean squared error cost, which was decided or use due to the nature of the problem where it is a regression. MSE is the sum of the square distances between our target variable and the predicted values, the cost function according to the number of Epochs is shown below:

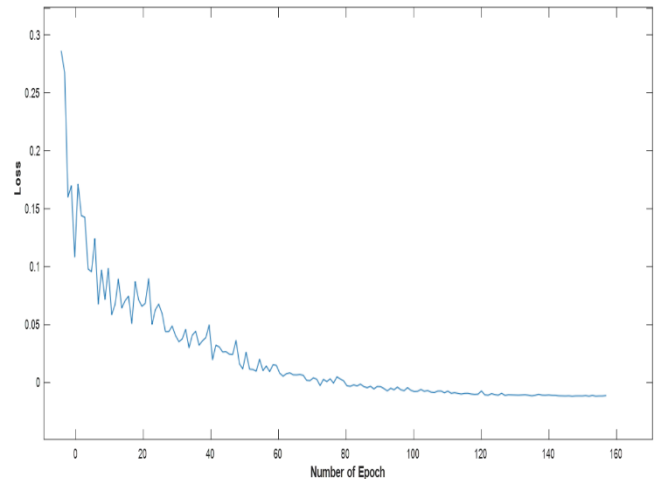


Figure VIII. Low MSE rate by increasing number of iterations.

The velocity planning approach is based on getting speed limits, limiting long-tied acceleration, and curving segments and side acceleration on curved segments to allow the UAV to fly with better control. By implementing robust path-planning

control, we are able to deliver precise path following regardless of uncertainties and variations in speed, on the other hand our trajectory resulting is 36.9m with 2.7 time second. In our work, we have made a speed profile that shows the changes in the speed while avoiding the obstacle in the environment. It can be seen from Table I the result of our work in comparison with the other related algorithms.

Table I. Result comparison with respect to trajectory, accuracy, and time.

Mode	Dijkstra	RRT	A*	Proposed
Resulting trajectory [m]	-	37.1015	-	36.9609
Accuracy	-	95%	-	98.6%
Time [s]	3.02996	6.9513	3.2781	2.7561

The idea behind the combination between the k-means and the ANN is to predict a suitable path has the ability to avoid the obstacle till the drone reach its goal, from the result above we can conclude that our work is better in part of length and speed. The problem with RRT considered as not smooth with highly random path so the path is not suitable. Comparing our work with the related shows that we have the ability to avoid the obstacle as fast as possible while if we just work with RRT algorithms we may have too many problems like, time speed, avoiding capabilities.

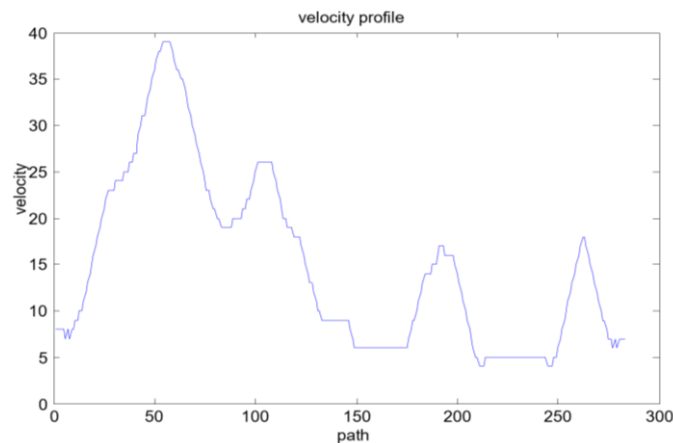


Figure IX. Speed profile path planning.

Fig. IX shows the performance of the drone moving from the start point trying to reach its goal. As we can see from the figure the speed is highest at the beginning of the movement that means the drone moving on the path after a while the speed will start decreasing because it's avoiding the obstacle, the speed still increasing and decreasing depending on the obstacle in the environment till the drone reaches its goal as safe as possible.

#### IV. Conclusions

In this paper, we presented an approach to guide a drone to fly through a 3D space area. The approach uses a K-means algorithm to segment the known area, a convolutional neural network to extract the features of the area segments, and ANN to classify these segments. Once the segments are defined and labeled, the drone can easily pass through the safe segments and avoid the unsafe segments, we evaluate our scheme using the MSE index and found that our approach has low MSE which decreases with each iteration of ANN training, also the

speed still increasing and decreasing depending on the obstacle in the environment till the drone reaches its goal.

In the future, we propose applying this scheme to a real-life drone with sensing devices to sense the surrounding area in real-time to guide the drone in uncharted areas.

#### References

- [1] Barraquand, Jerome, and Jean-Claude Latombe. "Robot motion planning: A distributed representation approach." *The International Journal of Robotics Research* 10.6 (1991): 628-649.
- [2] Mustafa mohammed, Oguz Ata, and Dogu Cagdas "Nonholonomic path planning for a mobile robot based on Voronoi and Q-Learning algorithm", 1st International Conference on Computing and Machine Intelligence (ICMI), pp.236-239, 2021.
- [3] Erdmann, Michael, and Tomas Lozano-Perez. "On multiple moving objects." *Algorithmica* 2.1 (1987): 477-521.
- [4] Karaman, Sertac, and Emilio Frazzoli. "Sampling-based algorithms for optimal motion planning." *The international journal of robotics research* 30.7 (2011): 846-894.
- [5] L. Zong, J. Luo, M. Wang, and J. Yuan, "Obstacle avoidance handling and mixed integer predictive control for space robots," *Adv. Sp. Res.*, vol. 61, no. 8, pp. 1997-2009, 2018, doi: 10.1016/j.asr.2018.01.025.
- [6] W. Zhao, H. Chu, M. Zhang, T. Sun, and L. Guo, "Flocking Control of Fixed-Wing UAVs with Cooperative Obstacle Avoidance Capability," *IEEE Access*, vol. 7, pp. 17798-17808, 2019, doi: 10.1109/ACCESS.2019.2895643.
- [7] J. Zhang, S. Zhang, and R. Gao, "Discrete-time predictive trajectory tracking control for nonholonomic mobile robots with obstacle avoidance," *Int. J. Adv. Robot. Syst.*, vol. 16, no. 5, pp. 1-11, 2019, doi: 10.1177/1729881419877316.
- [8] Y. Yu, X. Chen, Z. Lu, F. Li, and B. Zhang, "Obstacle avoidance behavior of swarm robots based on aggregation and disaggregation method," *Simulation*, vol. 93, no. 11, pp. 885-898, 2017, doi: 10.1177/0037549717711281.
- [9] P. Wu et al., "Autonomous obstacle avoidance of an unmanned surface vehicle based on cooperative manoeuvring," *Ind. Rob.*, vol. 44, no. 1, pp. 64-74, 2017, doi: 10.1108/IR-04-2016-0127.
- [10] P. Wang, S. Gao, L. Li, B. Sun, and S. Cheng, "Obstacle avoidance path planning design for autonomous driving vehicles based on an improved artificial potential field algorithm," *Energies*, vol. 12, no. 12, 2019, doi: 10.3390/en12122342.
- [11] L. Song, B. Y. Su, C. Z. Dong, D. W. Shen, E. Z. Xiang, and F. P. Mao, "A two-level dynamic obstacle avoidance algorithm for unmanned surface vehicles," *Ocean Eng.*, vol. 170, no. October, pp. 351-360, 2018, doi: 10.1016/j.oceaneng.2018.10.008.
- [12] J. Seo, Y. Kim, S. Kim, and A. Tsourdos, "Collision Avoidance Strategies for Unmanned Aerial Vehicles in Formation Flight," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 53, no. 6, pp. 2718-2734, 2017, doi: 10.1109/TAES.2017.2714898.
- [13] R. P. Padhy, S. K. Choudhury, P. K. Sa, and S. Bakshi, "Obstacle Avoidance for Unmanned Aerial Vehicles: Using Visual Features in Unknown Environments," *IEEE Consum. Electron. Mag.*, vol. 8, no. 3, pp. 74-80, 2019, doi: 10.1109/MCE.2019.2892280.
- [14] J. L. Mendoza-Soto, L. Alvarez-Icaza, and H. Rodríguez-Cortés, "Constrained generalized predictive control for obstacle avoidance in a quadcopter," *Robotica*, vol. 36, no. 9, pp. 1363-1385, 2018, doi: 10.1017/S026357471800036X.
- [15] X. Ma and A. Lee, "Self-self-adaptive obstacle avoidance fuzzy system of mobile robots," *J. Intell. Fuzzy Syst.*, vol. 35, no. 4, pp. 4399-4409, 2018, doi: 10.3233/JIFS-169759.
- [16] Z. Lin, L. Castano, E. Mortimer, and H. Xu, "Fast 3D Collision Avoidance Algorithm for Fixed Wing UAS," *J. Intell. Robot. Syst. Theory Appl.*, vol. 97, no. 3-4, pp. 577-604, 2020, doi: 10.1007/s10846-019-01037-7.
- [17] J. Li, J. Sun, and G. Chen, "A multi-switching tracking control scheme for autonomous mobile robot in unknown obstacle

- environments,” *Electron.*, vol. 9, no. 1, 2020, doi: 10.3390/electronics9010042.
- [18] L. Hwang, H. M. Wu, and J. Y. Lai, “On-Line Obstacle Detection, Avoidance, and Mapping of an Outdoor Quadrotor Using EKF-Based Fuzzy Tracking Incremental Control,” *IEEE Access*, vol. 7, pp. 160203–160216, 2019, doi: 10.1109/ACCESS.2019.2950324.
- [19] F. Belkhouche and B. Bendjilali, “Reactive path planning for 3-d autonomous vehicles,” *IEEE Trans. Control Syst. Technol.*, vol. 20, no. 1, pp. 249–256, 2012, doi: 10.1109/TCST.2011.2111372.
- [20] Zhang, Denggui, Yong Xu, and Xingting Yao. "An improved path planning algorithm for unmanned aerial vehicle based on rrt-connect." 2018 37th Chinese Control Conference (CCC). IEEE, 2018.
- [21] He, ZeFang, and Long Zhao. "The comparison of four UAV path planning algorithms based on geometry search algorithm." 2017 9th International Conference on Intelligent Human-Machine Systems and Cybernetics (IHMSC). Vol. 2. IEEE, 2017.

# Joint User Selection and Base Station Assignment Strategy in Smart Grid

Muhammad Fawad Khan

Dept. of Computer Science and Engineering  
Kyungpook National University  
Daegu, South Korea  
fawadkhan896@gmail.com

Muhammad Azam

School of Energy and Power Engineering  
Jiangsu University  
Zhenjiang, China  
azamlyh@gmail.com

Ashfaq Ahmed

Dept. of Electrical Engg. and CS  
Khalifa University  
Abu Dhabi, UAE  
ashfaq2419@gmail.com

Adnan Ahmed Khan

Department of Computer Science  
MCS, National University of Sciences  
and Technology  
Rawalpindi, Pakistan  
adnankhan@mcs.edu.pk

Muhammad Naeem

Department of Electrical Engineering  
COMSATS Univeristy Islamabad  
Wah Campus  
Wah, Pakistan  
muhammadnaeem@gmail.com

Imtiaz Alam

Department of Electrical Engineering  
Foundation University Islamabad  
Rawalpindi Campus  
Rawalpindi, Pakistan  
imtiazalam99@gmail.com

**Abstract**—Communication network is the most important part of emerging smart grid system. This information and communication network converts the traditional grid system to smart grids. The mere focus of this paper is to give the best combination of user and BSs that maximizes the overall throughput of the network. The traditional approach of user association comprises max-SINR (Signal to interference plus noise ratio) rule. According to which resources are allocated on the basis of max-SINR which is not suitable for heterogenous networks. In this work, overall performance of the network is optimized by using heuristic as well as optimal approach (fmincon). Proposed algorithm have the capability to find the best BS and user combination for maximization of overall throughput of the network. It also finds the maximum number of users that can be accommodated for a specific fixed number of BSs (Base Stations). Computational complexity has been remarkably reduced as compared to the existing work.

**Keywords**—smart grid, user association, optimization, base station assignment

## I. INTRODUCTION

Smart grids are the advanced form of conventional electrical grid systems where electricity is being supplied in a controlled manner [1]. There is a communication link between power generation, transmission and distribution that helps in improving the overall performance of grid systems. It is that communication or information exchange that converts the typical grid system to smarter one. In a brief context, smart grid is basically an automated electrical grid system that is more reliable, secure and efficient as compared to typical electrical grids. It supplies electricity according to consumer's demand in order to utilize the energy efficiently. As information and communication technology plays a pivotal role among the major stake holders of smart grid system. So, it can be termed as the backbone or heart of smart grid system. As shown in Fig 1, the scenario will be totally change by simply removing the block of information and communication

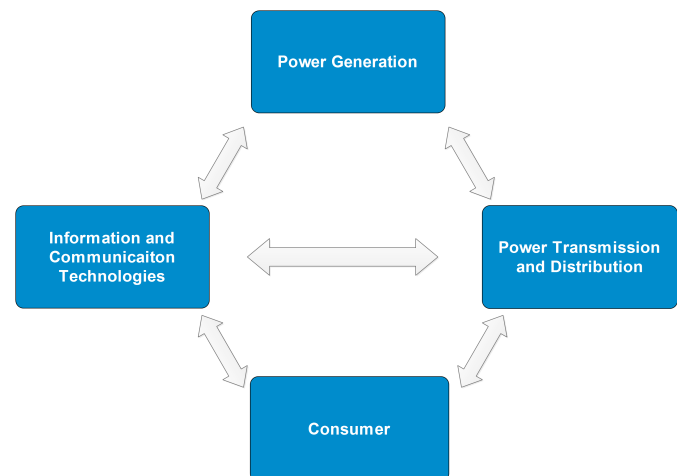


Fig. 1. Major stake holders of smart grid

technology and rest of the block will portray the true picture of conventional electrical grid systems. For a successful smart grid system, smart metering, energy management for both supplier and consumer as well as advanced communication methods and infrastructure are the core requirements [2], [3].

The emerging smart system merely rely on communication network. Usually wireless communication network is preferred as compared to wired link due to advantages like flexibility, ease of deployment, cutting cost and greater access [4]. Smart grid network comprises sensors, smart meters, a proper data management and monitoring system, various alarms, notifications, alerts and important indication mechanism. Keeping in mind the architecture of smart grid, communication network could be categorized into short and long range communications. Short range communication network is just like personal area network and is termed as Home Area Networks (HAN's).

It includes interfacing and communication between sensors, smart meters and devices deployed within the home premises. Bluetooth, Wifi, Zigbee, 6LoWPAN and Z-wave are some of the suitable communication technologies for HAN's.

Neighborhood Area Networks (NAN's) connects HAN's devices to the gateway network. NAN's are long range communication network. It is of worth importance in communication architecture of smart grid systems. There exists numerous options of communication technologies for NAN's i.e. satellite communication, WiMAX and cellular networks. Among these options cellular networks are preferred for NAN's due to high maturity of network, dedicated frequency band, high uplink (UL) and downlink (DL) data rates, lower latency, reliability, secure communication, ease of deployment and ubiquitous coverage [5].

Evolution of cellular network from 1G-4G and the exponential growth of devices, sensors and users are the true witness of drastic technological advancement in cellular network. Objectives of 5G targets 1000 times higher spectral efficiency, controlled privacy and reliable communication with 10 times lower energy consumption. Heterogenous networks, mm wave communication, D2D communication, massive MIMO, energy aware networks are considered as state of the art technologies for 5G networks. But still user association and BS assignment is of great importance in order to improving the overall performance of cellular technology for next generation network.

User association or BS assignment is the decision making for a particular user to get connect with a specific BS. This assignment or connection establishment is carried out using some specific decision variable targeting the optimization of the network. So, there is a requirement of an appropriate assignment strategy for both efficient utilization of resources and optimized performance of a particular communication network. This paper will focus on user selection and BS assignment strategy at the same time to get optimal results in order to meet the objectives.

## II. EXISTING WORK

Modern cellular networks are the most suitable technology for smart grid communications. For optimization of this communication network, researchers have discussed users and BS association problem extensively. In cellular architecture there are access nodes (BS) and a user must establish its connection with a specific access node (BS) to get the provided services. But [6] highlighted Coordinated Multipoint Technology (CoMP) where a single user can even establish connection with more than one BSs. New access nodes have been deployed in Het. Nets (Heterogenous Networks) in order to offload the traffic from macro BSs (Base stations). This deployment of pico/femto cell for handling more number of users at crowded areas is termed as cell splitting. Small cells (pico/femto) can hardly attract the users as they are low power cells as compared to macro BS. Transmit power has an important role for making a connection between user and the access node. Conventional approach of BS assignment strategy uses SINR (Signal to Interference Plus Noise Ratio ) as the decision parameter for

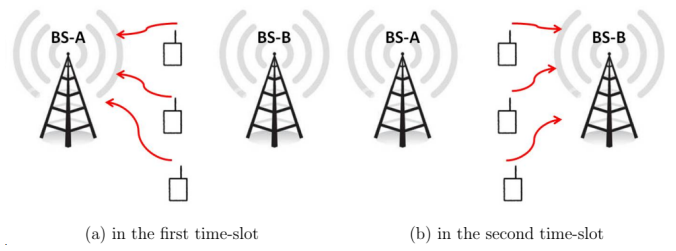


Fig. 2. Oscillatory behavior of greedy approach.

connecting the users to a particular BS. This conventional scheme is termed as max-SINR rule which failed in balancing the load between the macro and micro BSs. For this purpose [7] proposed a heuristic that adds a bias term in SINR of low power micro BS in order to push it to a comparable level to attract the users. [8]–[14] tried to minimize the interference by minimizing the total transmit power considering the minimum requirement for SNR constraint for every user. In existing literature, max-SINR rule has been discussed most of the time. But this typical approach has some serious issues regarding load balancing and transmit power of micro BSs. So, different researchers proposed numerous solutions for this problem. For example [15], targeted coordinate descent, dual coordinate descent as well as sub gradient method to get the association problem solved. Pricing variables are introduced in the objective function and their weightage actually helps in finding the optimal solution for the defined objective function. Some of the work of [15] have already been published in [16]. For increasing the coverage area and offloading traffic from high power macro BS [7] added a constant bias term to SINR. But it is difficult to specify the exact value for that optimal bias term. A greedy algorithm have been discussed in [17]–[19]. Drawback of BS assignment using greedy approach is that user will connect to a particular BS only considering the maximization of their own utility. Every time a user switch its BS on the basis of its own throughput in order to increase its own utility irrespective of the other parameters of network environment. It is difficult to control this type of mechanism. As shown in Fig. 2, each time user switch their BS for maximizing its own data rate which leads to a back and forth switching. It lies under the category of selfish algorithms leading to oscillatory behavior of the network. [20] proposed game theory for assignment of users and BSs. In [21], channel state information (CSI) has not been shared because of selfish reasons.

Rest of the paper has been organized in such a way that section III comprises system model and problem formulation while simulation results and discussion on the obtained results has been deeply discussed in section IV. Conclusion and future work has been highlighted in section V.

TABLE I  
NOTATIONS

Symbol	Definition
$N$	Total number of active users
$u_j$	User connected to each BS
$h_{ij}$	Channel between user $i$ and BS $j$
$b_{ij}$	Binary decision variable for association of $i_{th}$ user with $j_{th}$ BS
$P_j$	PSD (Power Spectral Density) of $j_{th}$ BS
$\gamma^2$	PSD of AWGN (Additive White Gaussian Noise)
$R_{ij}$	Data rate
$\Gamma$	SNR (Signal to Noise ratio) gap
$SINR$	Signal to interference plus noise ratio

### III. SYSTEM MODEL AND PROBLEM FORMULATION

The main focus of this work is on BS and user assignment strategy that yields the maximization of our objective. Choosing the cellular network as the communication network for smart grid systems, the target is to improve the overall throughput by selecting the best combinations of users and BSs.

Considering a downlink cellular network, let  $L$  be the number of BSs and total  $U$  active users are there within the area covered by the cellular network.  $i$  and  $j$  are used for indexing a particular user and BS respectively i.e.  $i \in [1, 2, 3, \dots, U]$  and  $j \in [1, 2, 3, \dots, L]$ .  $B$  is the total bandwidth (B.W) that is distributed among all the BSs. For making problem simpler, frequency flat power spectral density (PSD) levels and flat fading channels are assumed to make SINR values constant for all the frequencies. Channel between  $i_{th}$  user and  $j_{th}$  BS is denoted by  $h_{ij} \in \mathbb{C}$  and  $P_j$  is PSD level. Then SINR value for association of user  $i$  and BS  $j$  will be

$$SINR_{ij} = \frac{|h_{ij}|^2 P_j}{\sum_{j' \neq j} |h_{ij'}|^2 P_{j'} + \gamma^2} \quad (1)$$

$\gamma^2 =$  PSD of AWGN (Additive White Gaussian Noise).

In this paper, we adopt sum log utility maximization objective for optimizing the throughput by selecting the best combination of users and BSs. If  $u_j$  is the number of user associated with each  $j$  BS, then each user connected to that specific BS will share  $1/u_j$  of frequency resource. The data rate of user  $i$  connected with BS  $j$  is calculated by

$$R_{ij} = \frac{B}{u_j} \log \left( 1 + \frac{SINR_{ij}}{\Gamma} \right) \quad (2)$$

$\Gamma =$  SNR gap of user  $i$  that is determined by coding and modulation scheme.

$\Gamma$  is assumed to be same for every user.

Binary variable (0-1)  $b_{ij}$  is a decision variable used to determine whether user  $i$  is associated with BS  $j$  or not. Then objective function problem can be written as

$$f_o(\mathbf{b}, \mathbf{R}) = \sum_{i,j} b_{ij} \log(R_{ij}) \quad (3)$$

Introducing a new parameter termed as utility parameter  $a_{ij}$  as

$$a_{ij} = \log \left( B \log \left( 1 + \frac{SINR_{ij}}{\Gamma} \right) \right) \quad (4)$$

By substituting the value of  $a_{ij}$  into the objective function, our BS assignment problem becomes

$$\max_{\mathbf{b}, \mathbf{U}} \sum_{i,j} a_{ij} b_{ij} - \sum_j u_j \log(u_j)$$

subject to:

$$C1 : \sum_j b_{ij} = 1, \quad \forall i$$

$$C2 : \sum_i b_{ij} = u_j, \quad \forall j$$

$$C3 : \sum_j u_j = U$$

$$b_{ij} \in (0, 1), \quad \forall i, \forall j$$

C1 indicates one user can be connected to one BS at a time i.e. one-to-one connection between user and BS. C2 shows the total number of users connected to each BS denoted by  $u_j$ . C3 shows that all the users will be served.

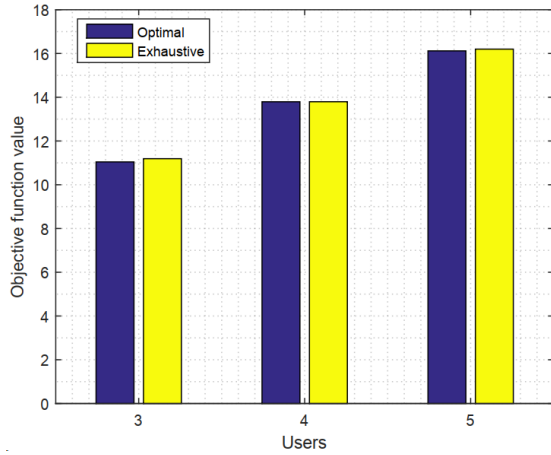
### IV. RESULTS AND DISCUSSION

In this section, the obtained result from exhaustive search and specified algorithm (fmincon) will be discussed extensively. Our objective function is a maximization function that depends on the user and BS assignment strategy.  $b_{ij}$  assigns a particular user to a BS and  $a_{ij}$  is the utility parameter introduced to get an estimate of maximum capacity (data rate) of the network.

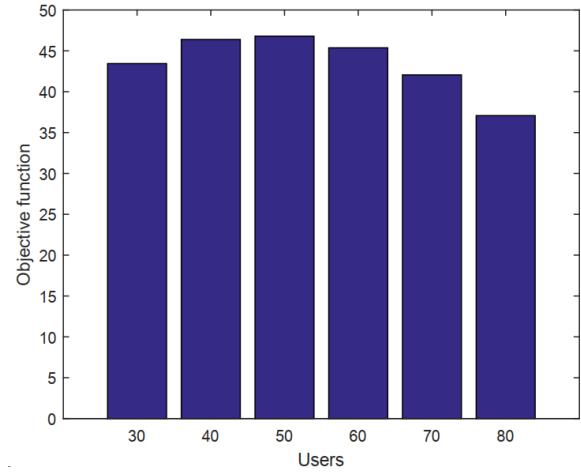
In Fig.3(a), exhaustive vs optimal algorithm output shows the value of objective function that is throughput in monetary units. There is a comparison in this figure between exhaustive and optimal search algorithm. It can be observed by the bar graph that optimal algorithm cannot give better results than the heuristic approach. Heuristic approach will always be better as compared to any algorithm. As indicated in graph, overall throughput of the network increases by increasing the number of users because utility increases in this way by fixing the number of BSs. If the number of BSs are fixed and we keep on increasing the users, then after specific number of users, the utility of the network start decreasing. This decrement indicates the point of saturation with specific number of BSs that can be observed in case of optimal algorithm.

Fig.3(b),(c),(d) shows the results of optimization algorithm i.e. (fmincon) used to improve the overall performance of the network with different combinations of users and BSs. For every set of users, the optimization algorithm calculates the objective function using the best assignment combination of users and BSs. There are three different graphs with unique throughput which indicates that different number of BSs are used. As indicated in Fig.3(b), 2 BSs are fixed for this case. When we keep on increasing the number of users with fixed

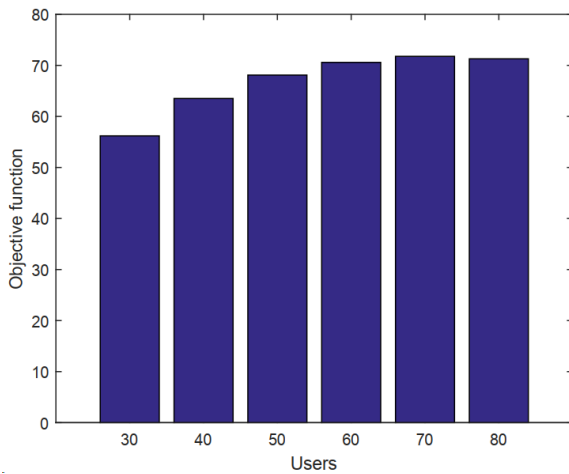




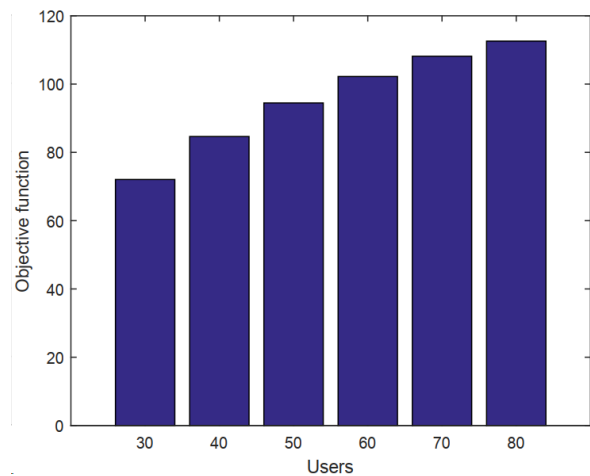
(a) Exhaustive vs Optimal algorithm



(b) Max achievable throughput with 2 BSs



(c) Max achievable throughput with 3 BSs



(d) Max achievable throughput with 5 BSs

Fig. 3. (a) Throughput performance comparison of exhaustive vs optimal algorithm in monetary units. (b), (c), (d) indicates the max achievable throughput and the saturation point for the accommodation of users with specific number of base stations.

number of BSs, the network throughput increases till 50 users. After 50 users, the throughput decreases which indicates the overloaded traffic on 2 BSs. So, for best utilization, 2 BSs can accommodate max 50 users in order to get maximum overall throughput of the network. If we want to accommodate more users then number of BSs must be increased for better services. In the same way, if we increase number of BSs to 3 and 5 as indicated in Fig.3(c) and (d), we can provide best services to 70 and even more number of users taking the best assignment strategy using `fmincon`. In a nutshell, our algorithm selects users and BSs assignment jointly targeting the maximum throughput of the network.

## V. CONCLUSION AND FUTURE WORK

This paper jointly optimizes user and BS association by selecting the best combination of user and BS that maximizes

the overall throughput of the network. Implementing the heuristic approach gives the maximum results as compared to optimal algorithm. For any specific number of BSs, saturation point of user accommodation has been identified. This saturation point is directly linked with the utility of the objective function and identify that maximum number of users that can be served using that fixed amount of BSs. It is clearly observed from the results that by increasing the number users increases the overall throughput till saturation point and after this point any increase in user leads to decrement in the utility of the network. The optimization algorithm (`fmincon`) yield the best combination of user and BS that is then used for getting maximum throughput up till certain saturation level. So, keeping the saturation point in mind, the performance of the network has been optimized using optimal algorithm.

Jointly optimizing the power as well as user association could be a promoting extension of this work. Beamforming and power control are still two important aspects of a communication network that can make it more suitable and efficient for emerging smart grid communication network.

## REFERENCES

- [1] P. Siano, "Demand response and smart grids—a survey," *Renewable and sustainable energy reviews*, vol. 30, pp. 461–478, 2014.
- [2] Y. Kabalci, "A survey on smart metering and smart grid communication," *Renewable and Sustainable Energy Reviews*, vol. 57, pp. 302–318, 2016.
- [3] Y. Yan, Y. Qian, H. Sharif, and D. Tipper, "A survey on smart grid communication infrastructures: Motivations, requirements and challenges," *IEEE communications surveys & tutorials*, vol. 15, no. 1, pp. 5–20, 2013.
- [4] M. Rafiei, S. M. Elmi, and A. Zare, "Wireless communication protocols for smart metering applications in power distribution networks," in *Electrical Power Distribution Networks (EPDC), 2012 Proceedings of 17th Conference on*. IEEE, 2012, pp. 1–5.
- [5] C. Kalalas, L. Thrybom, and J. Alonso-Zarate, "Cellular communications for smart grid neighborhood area networks: A survey," *IEEE Access*, vol. 4, pp. 1469–1493, 2016.
- [6] E. U. T. R. Access, "Further advancements for e-utra physical layer aspects," *3GPP Technical Specification TR*, vol. 36, p. V2, 2010.
- [7] I. Guvenc, M.-R. Jeong, I. Demirdogen, B. Kecioglu, and F. Watanabe, "Range expansion and inter-cell interference coordination (icic) for picocell networks," in *Vehicular Technology Conference (VTC Fall), 2011 IEEE*. IEEE, 2011, pp. 1–6.
- [8] L. Smolyar, I. Bergel, and H. Messer, "Unified approach to joint power allocation and base assignment in nonorthogonal networks," *IEEE Transactions on Vehicular Technology*, vol. 58, no. 8, pp. 4576–4586, 2009.
- [9] J. T. Wang, "Sinr feedback-based integrated base-station assignment, diversity, and power control for wireless networks," *IEEE Transactions on Vehicular Technology*, vol. 59, no. 1, pp. 473–484, 2010.
- [10] H. Pennanen, A. Tölli, and M. Latva-aho, "Decentralized base station assignment in combination with downlink beamforming," in *Signal Processing Advances in Wireless Communications (SPAWC), 2010 IEEE Eleventh International Workshop on*. IEEE, 2010, pp. 1–5.
- [11] D. H. Nguyen and T. Le-Ngoc, "Joint beamforming design and base-station assignment in a coordinated multicell system," *Iet Communications*, vol. 7, no. 10, pp. 942–949, 2013.
- [12] V. N. Ha and L. B. Le, "Distributed base station association and power control for heterogeneous cellular networks," *IEEE Transactions on Vehicular Technology*, vol. 63, no. 1, pp. 282–296, 2014.
- [13] R. D. Yates and C.-Y. Huang, "Integrated power control and base station assignment," *IEEE Transactions on vehicular Technology*, vol. 44, no. 3, pp. 638–644, 1995.
- [14] S. V. Hanly, "An algorithm for combined cell-site selection and power control to maximize cellular spread spectrum capacity," *IEEE Journal on selected areas in communications*, vol. 13, no. 7, pp. 1332–1340, 1995.
- [15] K. Shen and W. Yu, "Distributed pricing-based user association for downlink heterogeneous cellular networks," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 6, pp. 1100–1113, 2014.
- [16] —, "Downlink cell association optimization for heterogeneous networks via dual coordinate descent," in *Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on*. IEEE, 2013, pp. 4779–4783.
- [17] R. Madan, J. Borran, A. Sampath, N. Bhushan, A. Khandekar, and T. Ji, "Cell association and interference coordination in heterogeneous lte-a cellular networks," *IEEE Journal on selected areas in communications*, vol. 28, no. 9, pp. 1479–1489, 2010.
- [18] T. B. L. E. L. R. Ramjee, "Generalized proportional fair scheduling in third generation wireless data networks," in *IEEE INFOCOM, 2006*, pp. 1–12.
- [19] K. Son, S. Chong, and G. De Veciana, "Dynamic association for load balancing and interference avoidance in multi-cell networks," *IEEE Transactions on Wireless Communications*, vol. 8, no. 7, 2009.
- [20] L. Jiang, S. Parekh, and J. Walrand, "Base station association game in multi-cell wireless networks (special paper)," in *Wireless Communications and Networking Conference, 2008. WCNC 2008. IEEE*. IEEE, 2008, pp. 1616–1621.
- [21] M. Hong and A. Garcia, "Mechanism design for base station association and resource allocation in downlink ofdma network," *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 11, pp. 2238–2250, 2012.

# Comparison of flowdrill and conventional drilling methods in thin-walled materials

Mert Safak TUNALIOGLU  
Department of Mechanical Engineering  
Hitit University  
Corum, Turkey  
mertstunalioglu@hitit.edu.tr

**Abstract**—In this study, drilling and tapping processes with flowdrill method in thin-walled materials were investigated. Holes were drilled in AISI 304 stainless steel and St 37-2 materials with a wall thickness of 1.5 mm with different diameters of flowdrill and conventional drilling methods. Taps were drilled into the drilled holes with flowtap and traditional tapping methods. Obtained results were compared in terms of the bushing heights, the strength of the joints, micro hardness around the holes, and observation of the capillary crack formations around the holes. When the flowdrill method is used in materials with thin wall thickness, 4-5 times the bushing height and as a result approximately 3.5-4 times the joint strength compared to the conventional drilling method. While capillary cracks were observed in large hole diameters in the conventional drilling process, no capillary cracks were found for each hole diameter tested in the flowdrill process.

**Keywords**—flowdrill, flowtap, conventional drilling, conventional tapping

## I. Introduction

In recent years, when automation has come into prominence, the most important rule is to process materials in the most efficient way in the shortest time with appropriate production methods. Materials with thin wall thickness, the pitch lengths are insufficient for tapping in drilling and a weak connection is provided. Besides, capillary cracks form around the hole during the drilling process due to the wear of the drill tip and this decreases the strength of the materials. There are several alternative methods to strengthen the joint such as welding nuts, tapping rivets, bonding, and using blind bolts. These methods provide an increase in strength in the connection, but also increase both cost and production times.

The flow drilling process is forming holes on parts by the effect of the contact of flow drill tip with the material at high speed. Sufficient pitch length required for the screw joint is easily obtained since there is no chip removal during the drilling process (Figure I). Due to this advantage, there is no need to weld a nut on the sheet metal part. During the flow drilling process, crack formation risk also is prevented since high temperature occurs around the hole. In the flow tapping process, warping and cracking risk is also prevented.

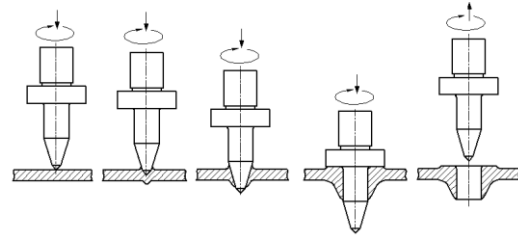


Figure I. Stages of flowdrill process [1]

The main purpose of the studies about drilling holes to thin-walled materials in the literature is increasing the strength of the joint. Especially in the studies conducted oriented to construction sector, applications such as welding or bonding hollow bolt or nutted rivet to the holes drilled on circular or square profile materials having different properties filled with concrete in T-type joints were made to increase the strength of the joint, and the increase of the joint strengths depending on the hole diameters and material thicknesses were investigated experimentally and theoretically [2-5]. Some researchers have shown that the strength of the connection is further increased with the extended hollow bolt, slip-critical blind bolt and tube bolts they have developed [6-9]. Researchers working in the flowdrill method, which was developed as an alternative to drilling, conducted experimental and theoretical studies on tool wear and drilling parameters with spinning (friction angle, tool advance, tool angle, spindle speed, tool rotation speed, thrust and axial force). they have done. [10-14]. In this study, AISI 304 stainless steel and St 37-2 materials having 1.5 mm thickness were flowdrilled and flowtapped in various diameters. Same processes were repeated with conventional drilling method. Obtained results were compared in terms of the bushing heights, the strength of the joints, micro hardness around the holes, and observation of the capillary crack formations around the holes.

## II. Theory and Method

In this study, drilling and tapping processes were performed on AISI 304 stainless steel and St 37-2 materials with a wall thickness of 1.5 mm with flowdrill and flowtap methods, respectively. Long necked flowdrill tips with diameters of  $\text{Ø}4.6$  mm,  $\text{Ø}5.5$  mm and  $\text{Ø}7.3$  mm were used for drilling with the flowdrill method (Figure II). Scrubbing type M5, M6 and M8 taps were used to flowtap method (Figure III). Conventional drilling and tapping processes were applied to the same materials. The mechanical and chemical properties of AISI 304 stainless steel and St 37-2 materials used in the study are shown in Tables I and II.

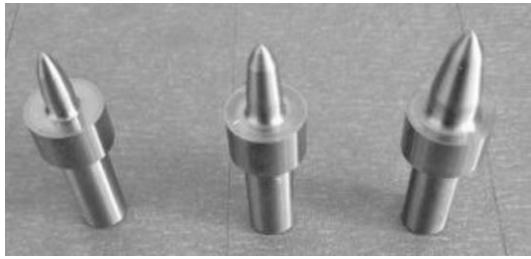


Figure II. Flowdrill tips

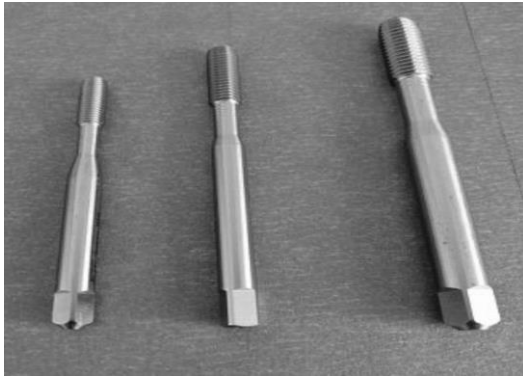


Figure III. Flowtap tips

Table I. Technical properties of the materials

Properties name	AISI 304	St 37-2
Hardness (HB)	201	120
Ultimate Tensile Strength (MPa)	505	360
Ultimate Yield Strength (MPa)	215	215
Elongation at Break (%)	70	24
Modulus of Elasticity (GPa)	193	212
Poisson Ratio	0.29	0.28
Shear Modulus (GPa)	86	82
Density (g/cc)	8.00	7.96
Specific Heat Capacity (J/g°C)	0.5	0.42
Thermal Conductivity (W/mK)	16.2	5.7
Melting Point (°C)	1455	1480

Table II. Chemical properties of the materials

Component Element Name	AISI 304	St 37-2
Carbon (C)	max. 0.08	0.20
Chrome (Cr)	18-20	-
Iron (Fe)	66.35-74	75
Manganese (Mn)	≤ 2	-
Phosphorous (P)	≤ 0.045	0.05
Sulphur (S)	≤ 0.03	0.05
Silicon (Si)	≤ 1	15-35
Nickel (Ni)	8-10.5	0.09

In the flowdrill method, the feed rate was taken as 125 mm/d for each test. Holes in Ø4.6 mm, Ø5.5 mm and Ø7.3 mm diameters at 2000 rpm were drilled into the test samples with a flowdrill tip, M5, M6 and M8 threads were drilled, respectively. In the conventional drilling method, chips were removed from the samples with a drill bit and Ø4.2 mm, Ø5 mm and Ø6.8 mm diameter holes were drilled and M5, M6 and M8 were drawn respectively. The test plan applied in the study is shown in Table II. Both methods were compared in

terms of bushing heights, clamping force, hardness and crack analyses.

Table III. Test plan

Material	Thickness (mm)	Hole & tapping dia. (mm)	Spindle speed (rpm)	Method
AISI 304	1,5	Ø4,6-M5	2000	Flowdrill & Flowtap
		Ø5,5-M6		
		Ø7,3-M8		
St 37-2	1,5	Ø4,6-M5	2000	Flowdrill & Flowtap
		Ø5,5-M6		
		Ø7,3-M8		
AISI 304	1,5	Ø4,6-M5	2000	Conventional Drilling & Tapping
		Ø5,5-M6		
		Ø7,3-M8		
St 37-2	1,5	Ø4,6-M5	2000	Conventional Drilling & Tapping
		Ø5,5-M6		
		Ø7,3-M8		

### A. Bushing height measurements of flowdrilling and conventional drilling methods

In accordance with Table III, the bushing heights of the test samples drilled with the flowdrill and conventional drilling methods were measured with the help of a digital caliper (Figure IV). For the accuracy of each test, measurements were made on 3 samples from the same group, and the average of these measurements was taken. The bushing heights of the samples drilled with the flowdrill and conventional drilling methods were compared.



Figure IV. Bushing height measurement with digital caliper

### B. Clamping strengths measurements of flowdrilling and conventional drilling methods

To examine the strength of thin-walled materials, clamping tests were performed on the test specimens. Since the material used in the study is thin-walled, a special apparatus was made to prevent deformation in the clamping tests (Figure V). The materials drilled and tapped with both methods were fastened with M5, M6 and M8 bolts specified in Table III and their clamping strengths were determined.



Figure V. Clamping strength test apparatus

### C. Hardness measurements of flowdrilling and conventional drilling methods

During the drilling process, the drilling tip is contacted with a force on the material at high speeds and high temperatures occur due to the friction that occurs. These temperatures affect the hardness of the hole region. 3 measurements were made for each sample from the region close to the holes drilled with the flowdrill and tradition drilling methods.

### D. Penetrant test measurements of flowdrilling and conventional drilling methods

Penetrant testing is a non-destructive testing method used to detect surface defects (Figure VI). Since the wall thickness of the test samples are thin, the capillary cracks that may occur around the hole as a result of drilling were examined with the help of penetrant tests. Penetrant tests were applied to the outer parts of the holes and the root areas of the bushes, which were applied by the flowdrill and conventional drilling method. The capillary cracks in the samples were determined by applying spray to these areas.



Figure VI. Penetrant test samples

## III. Test Results

### A. Comparison of flowdrilling and conventional drilling methods in terms of bushing height

In this study materials drilled in 1.5 mm wall thickness AISI 304 stainless steel and St 37-2 at different diameters using flowdrill and conventional drilling methods, and the graphs of the bushing heights measured with the digital caliper are shown comparatively in Figures VII and VIII.

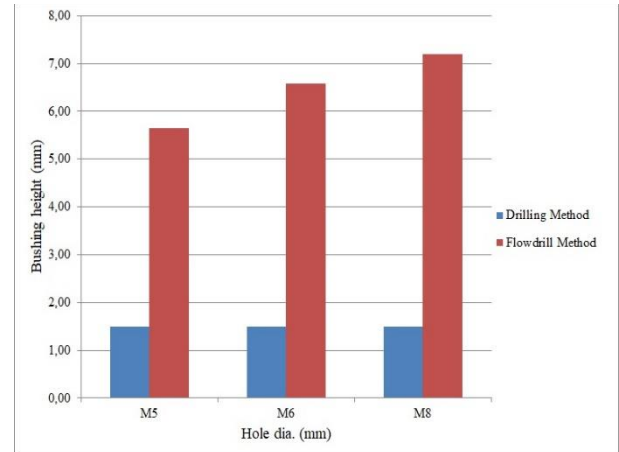


Figure VII. AISI 304 stainless steel bushing heights test results

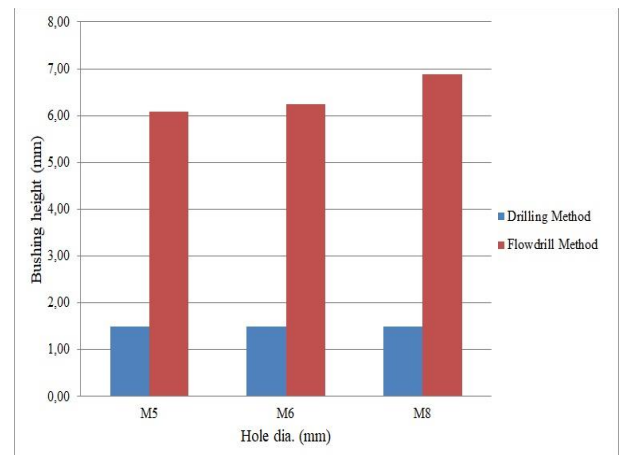


Figure VIII. St 37-2 bushing heights test results

The bushing heights formed which are equal to the wall thickness of the test specimens are 1.5 mm when both materials are drilled with the conventional drilling method (Figure VII-VIII). According to Figures VII and VIII when the hole diameters increase in AISI 304 stainless steel and St 37-2 materials, the amount of drilled material and the bushing heights increase.

When the flowdrill and conventional drilling methods are compared, the bushing heights in AISI 304 stainless steel material increased 3.8 times for M5, 4.4 times for M6 and 4.8 times for M8, respectively (Figure VII). In the St 37-2 material, the bushing heights increased 4.1 times for M5, 4.2 times for M6 and 4.6 times for M8, respectively (Figure VIII).

As a result of the tests, the bushing heights of the holes drilled with the flowdrill method increased by 4-5 times compared to the conventional drilling method. With this increase comes the possibility of tapping more of the test specimens with a larger amount of tap. It is predicted that the connection will be strengthened with the increase in the height of the bushes in these ratios and with more threading in the parts.

### B. Comparison of flowdrilling and conventional drilling methods in terms of clamping strengths

To measure the strength of thin-walled connections, clamping tests were carried out on materials who drilled by flowdrill and conventional drilling methods. Clamping tests performed at 1 mm/min. a constant speed.

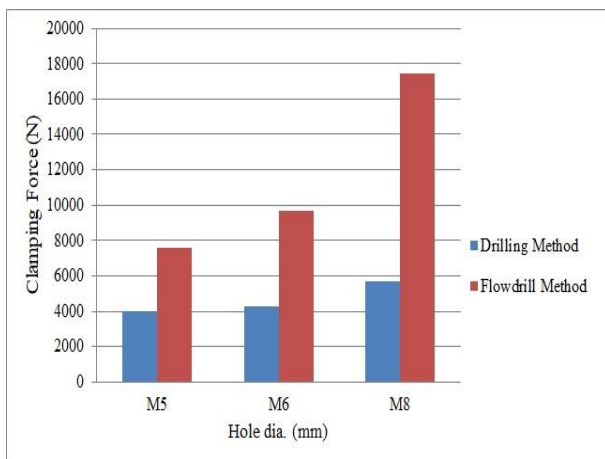


Figure IX. AISI 304 stainless steel clamping strengths test results

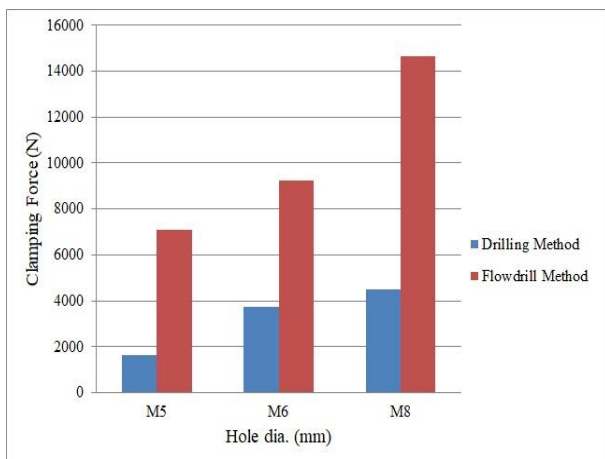


Figure X. St 37-2 clamping strengths test results

For both materials, the clamping strengths increased as the hole diameters increased (Figure IX-X). When the hole diameter increase, more material is plastered and the bushing height increases. As the height of the bushes increases, more threads are formed and these threads resist greater clamping strengths.

The clamping strength of AISI 304 stainless steel material increased 1.98 times for M5, 2.33 times for M6 and 3.02 for M8, respectively (Figure IX). In St37-2 material, the clamping strengths increased 4 times for M5, 2.35 times for M6 and 3.45 for M8, respectively (Figure X). When both materials were examined, the holes drilled by the flowdrill

method, the peel strength increases approximately 2-3 times compared to the conventional drilling method.

### C. Comparison of flowdrilling and conventional drilling methods in terms of part hardness

During drilling with flowdrill method, the friction of the flowdrill tip and the part at high speeds causes a temperature increase of approximately 300-400 °C around the holes. Hardness tests were performed for both materials in terms of hardness changes around the hole with sudden heating and cooling.

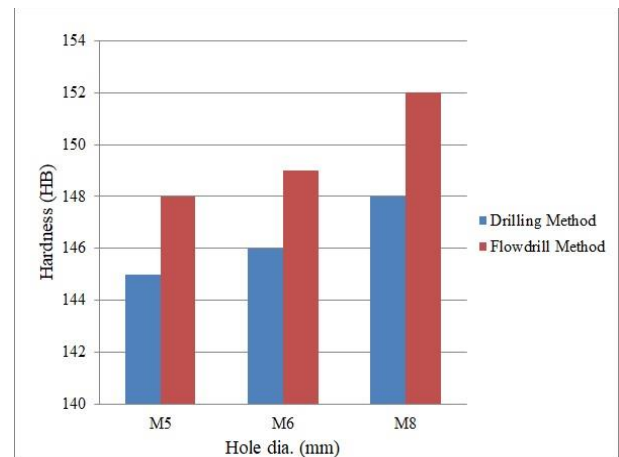


Figure XI. AISI 304 stainless steel clamping hardness test results

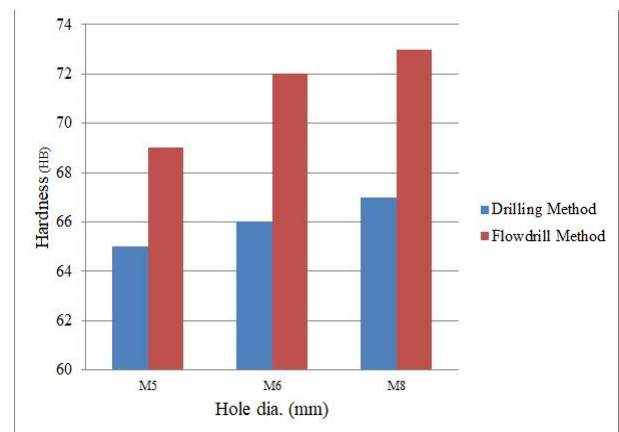


Figure XII. St 37-2 hardness test results

In Figure IX-X, the hardness values of both materials increased as the hole diameter increased. When the hardness values of AISI 304 stainless steel flowdrill and drilling methods are compared, an increase of approximately 4% is observed in AISI 304 stainless steel and approximately 6% in St37-2.

### D. Comparison of flowdrilling and conventional drilling methods in terms of penetrant test results

The results of the penetrant tests applied to the outer and the root region of the holes that were subjected to drilling and tapping processes with the flowdrill and conventional drilling methods are given in Tables IV and V comparatively.

Table. IV AISI 304 stainless steel and St 37-2 flowdrill method penetrant test results

AISI 304 stainless steel			St 37-2		
Hole dia. (mm)	Spindle speed (rpm)	Test result	Hole dia. (mm)	Spindle speed (rpm)	Test result
M5	2000	No crack	M5	2000	No crack
M5	2000	No crack	M5	2000	No crack
M5	2000	No crack	M5	2000	No crack
M6	2000	No crack	M6	2000	No crack
M6	2000	No crack	M6	2000	No crack
M6	2000	No crack	M6	2000	No crack
M8	2000	No crack	M8	2000	No crack
M8	2000	No crack	M8	2000	No crack
M8	2000	No crack	M8	2000	No crack

Table V. AISI 304 stainless steel and St 37-2 conventional drilling method penetrant test results

AISI 304 stainless steel			St 37-2		
Hole dia.	Spindle speed (rpm)	Test result	Hole dia.	Spindle speed (rpm)	Test result
M5	2000	No crack	M5	2000	No crack
M6	2000	No crack	M6	2000	Crack occurs
M8	2000	Crack occurs	M8	2000	Crack occurs

As a result of the penetrant tests applied to the samples drilled and tapped at different hole diameters with the flowdrill method, no capillary cracks were found around or in the root regions of the holes in both material (Table IV). No cracks were found around the M5 and M6 holes in the AISI 304 stainless steel and St37-2 materials drilled and tapped with the conventional drilling method. In both materials, capillary cracks were found in the M8 hole (Table V). In the conventional drilling method, crack formation is observed when the hole diameter exceeds certain dimensions. The formation of crack causes loss of strength in the materials.

#### IV. Conclusion

In this study the flowdrill method was examined to increase the strength in thin-walled materials and the results were compared with the conventional drilling method.

AISI 304 stainless steel and St37-2 materials with a wall thickness of 1.5 mm are approximately 4-5 times more advantageous in terms of the bushes heights that occur when the holes are drilled with flowdrill method. Holes tapped with the flowdrill method are 3.5-4 times more durable than conventional drilling. The walls of the holes drilled with the flowdrill method is approximately 4-5% harder than the conventional drilling method. While there were no capillary cracks around the holes drilled with the flowdrill method, the formation of capillary cracks began to be observed as the hole diameter increased in the materials drilled with the conventional drilling method.

Considering all these results, the flowdrill method is a suitable technique especially for thin-walled materials. With this method, fast and durable holes are drilled. It is more economical than other connection methods (blind bolts,

welding, rivets nuts, bonding). After the drilling process, the tapping can be opened easily, eliminating the risk of warping in the material.

#### References

- [1]. Nilufer Dogru, "Investigation of Machining Characteristics of Friction Drilling on AISI 1010 steel" Firat University, Turkey, 2010.
- [2]. John E. France, J..Buick Davison, Patrick A. Kirby, "Strength and rotational response of moment connections to tubular columns using flowdrill connectors", J. Constr. Steel Res. 50, 1–14, 1999.
- [3]. Johan Kolsto Sonstabo, David Morin, Magnus Langseth, "Macroscopic Modelling of Flow-drill Screw Connections in Thin-walled Aluminium Structures", Thin-Walled Structures. 105, 185-206, 2016.
- [4]. Johan Kolsto Sonstabo, Peter Henrick Holmstrom, David Morin, "Macroscopic Strength and Failure Properties of Flow-drill Screw Connections", Journal of Materials Processing Technology, 222, 1-12, 2014.
- [5]. Jingpeng Wang, Liping Chen, "Experimental Investigation of Extended end Plate Joint to Concrete-filleed Stell Tubular Columns", Journal of Constructional Steel Research, 79, 56-70, 2012.
- [6]. Walid Tizani, Ali Al-Mughairi, John S. Owen, Theodoros Pitrakos, "Rotational stiffness of a blind-bolted connection to concrete-filled tubes using modified Hollo-bolt", Journal of Constructional Steel Research, 80, 317–331, 2013.
- [7]. Wei Wang, Ling Li, Dabiao Chen, "Progressive collapse behaviour of endplate connections to cold-formed tubular column with novel Slip-Critical Blind Bolts", Thin-Walled Structures, 131, 404–416, 2018.
- [8]. Mahdi Zeinizadeh Jeedi, N. H. Ramli Sulong, "Pull-out performance of a novel anchor blind bolt (TubeBolt) for beam to concrete-filled tubular (CFT) column bolted connections", Thin-Walled Structures 124, 402–414, 2018.
- [9]. Wei Wang, Mingxiao Li, Yiyi Chen, Xiaogang Jian, "Cyclic behavior of endplate connections to tubular columns with novel slip-critical blind bolts", Engineering Structures, 148, 949–962, 2017.
- [10]. Cebeli Ozek, Zulkuf Demir, "Investigate the Surface Roughness and Bushing Shape in Friction Drilling Of A7075-T651 and St 37 Steel", TEM Journal, 2;2, 170-180, 2013.
- [11]. Zulkuf Demir, "Investigation of the Fluctuation Size in Thrust Force and Chip Morphology in Drilling, Celal Bayar University Journal of Science", 14,4, 385-397, 2018.
- [12]. Kuan-Yu Sua, Torgeir Welo, Jyhwen Wang, "Improving Friction Drilling and Joining through Controlled Material Flow", Procedia Manufacturing, 26, 663-670, 2018.
- [13]. Sara A. El-Bahloul, Hazem E. El-Shourbagy, Ahmed M. El-Bahloul, Tawfik T. El-Mindany, "Experimental and Thermo-Mechanical Modeling Optimization of Thermal Friction Drilling for AISI 304 Stainless steel", CIRP Journal of Manufacturing Science and Technology, 20, 84–92, 2018.
- [14]. N. Rayesh Jeudoss Haynes, R. Kumar, "Simulation on friction drilling process of Cu2C", Materials Today: Proceedings, 5, 27161–27165, 2018.

# Classification of Animal Faces Using a Novel DAG-CNN Architecture

Shahram Taheri  
Department of Computer Engineering  
Antalya B ilimUniversity  
Antalya, Turkey  
shahram.taheri@antalya.edu.tr

Zahra Golrizkhatami  
Department of Computer Engineering  
Antalya B ilimUniversity  
Antalya, Turkey  
z.golrizkhatami@antalya.edu.tr

Önsen Toygar  
Department of Computer Engineering  
Eastern Mediterranean University  
Famagusta, North Cyprus, Turkey  
onsen.toygar@emu.edu.tr

**Abstract**—Manual classification and recognition of animals in wild life images and footage is a tiring and extremely challenging process. Therefore, automatic systems that are developed by computer vision approaches for the classification of animals are suggested. In this research work, we present a new architecture of a non-linear deep learning structure namely Directed Acyclic Graph Convolutional Neural Networks (DAG-CNNs) for animal classification. This system applies several Convolutional Neural Network (CNN) layers' learned features and fuses them for the final decision making. For this purpose, popular and publicly available CNN architecture, namely VGG-16 is selected and considered as the underlying backbone structure of the proposed system and several new branches are added to it. The proposed system automatically performs the multi-stage feature extraction and combines multiple classifiers decisions in score-level fusion manner. Experiments on the open access animal database prove the capability and efficiency of this novel method.

**Keywords** convolutional neural network, score-level fusion, directed acyclic graph, animal classification

## I. Introduction

Animal classification is a subdomain of visual object categorization which has been reported in a few studies. Detection and classification of animals can be used in various applications of computer vision such as animal-vehicle accident prevention, animal trace facility, identification, antitheft of animals in zoo and content-based image retrieval. Animals are considered as one of the most difficult objects in object detection applications [1]. There are various reasons such as the point that most of the animals are able to self-mask and usually they come out in complicated scene with varying illumination, viewpoints and scales. Animal pictures captured in the wild may have complex backgrounds, various postures and diverse illuminations. Another difficulty in the task of animal classification is that the available datasets contain a limited number of animal classes. Animals are one of the object classes used in the state-of-the-art for object recognition. The performance of object recognition system strictly relies on how well the object representation and characterization are done.

Although various human face recognition methods have been proposed in literature, they are not fully appropriate for animal face classification with high range of intra-class variations and inter-class similarities. In this respect, several object recognition algorithms have been applied on animals' images with the aim of extracting hand-crafted features such as texture and shape. The major drawback of these methods is that, they are entirely problem dependent. Consequently, the productivity of these algorithms is problematic. Constructing the complex feature vector of various hand-crafted features which has high dimension will take extensive time which may not be efficient. Animal classification and recognition can also

be valuable in expert systems to determine the wild animals' migration corridors. Object characterization can be obtained by applying visual descriptors, shape descriptors or texture representation. Deep learning and specifically Convolutional Neural Network (CNN) approaches have been successfully employed in recent studies of object classification and recognition tasks [2-5] and shown to have salient performance comparing to state-of-the-art. CNN is an end-to-end system which is capable of extracting relevant features and also integrate the feature extraction and classification phases of classical machine learning systems. To train a CNN, presence of a huge set of samples is required. CNN extracts the discrimination features automatically which in many cases these features are shown to be superior to hand-crafted features such as HOG, LBP, or SURF [6]. Generally CNN architectures consists of various numbers of basic elements namely convolutional, Max-pooling and ReLU, followed by the multi-layer perceptron neural network. In our previous study [7], in order to take the advantage of both approaches, two different classifier systems were trained. One to employ the CNN features, and the other system had been trained based on hand-crafted features, such as appearance-based and shape-based features. Afterwards the outcomes were fused and the final results were obtained based on this fusion.

In this research, we develop a new CNN architecture, namely Directed Acyclic Graph Convolutional Neural Networks (DAG-CNNs) for animal classification. The proposed method exerts the multi-stage features from different CNN layers. DAG-CNN integrates the feature extraction and classification phases of CNN into a single automated learning process. Furthermore, the proposed system employs multi-stage features and carry out the score-level fusion of multiple classifiers automatically. Hence, the proposed system revokes the necessity of extracting the hand crafted feature.

The contributions of our work are as follows:

- The proposed system combines two main stages of conventional machine learning procedure, namely feature extraction and classification stages, into a complete automatic learning scheme for animal classification.
- Unlike the classic CNN which uses only the features of the last layer, the proposed system employs the multi stage learned features from mid-layers as well.
- The proposed system aggregates the results of different classifiers and automatically combines them by the score-level fusion approach and positively it negates the use of feature-level fusion.

The proposed architecture has been tested over LHI-animal-faces database which contains 20 different categories. 19 classes of various animal types and one class of human faces. The differentiation of these categories is very challenging due



to their evolutionary relationship and shared parts. Additionally, significant within-class variations such as rotation, flip transforms, posture variation and sub-types exist in the face categories. In order to make a comparison and following the state-of-the-art methods, in our experiments, 30 samples of each class were selected for training and the remaining samples for testing. The obtained results confirm that the usage of multi-stage features from different layers of a CNN remarkably improves the classification accuracy.

The rest of this paper is organized follows. Related works are reviewed in Section 2 and overview of DAG-CNN architecture is given in Section 3. Proposed method and experimental settings and the obtained results are presented in Section 4. Finally, the conclusion is stated in the last section.

## II. Related works

One of the earliest studies on animal recognition was performed by Schmid et al. [8]. They built an image retrieval system for 4 classes of animals by applying Gabor filters. Later, Ramanan et al. proposed systems for detecting animals in video frames by utilizing the texture and shape-based features [9-10]. In [11], the authors introduced a method for animal images search engine. An attempt for classifying marine species is performed by Cao et al. [12]. For this purpose, they fused CNN and local feature descriptors and achieved better result than individual system. Peng et al in [13] introduced a deep learning algorithm in order to classify animal face images and investigated its performance on LHI-Animal-Faces dataset. They proposed a deep boosting framework based on layer-by-layer joint feature boosting and dictionary learning. In each layer, they utilized a collection of filters by fusing the early layer filters.

Afkham et al. [14] collected a new dataset including realistic images of 13 various species with complex in wild background. They introduced a system for visual object classification based on combining the textural features in samples and their background. An approach for classifying 25 categories of animals is reported in [15]. The animal images are segmented and partitioned into small patches and then some colour-related features are extracted from each of them. These features are fed into probabilistic neural network for classification.

In [16], the authors suggested two systems for classifying 20 classes of animals by using Gabor features and K-mean algorithm. Matuska et al. [17] developed a system which detects and classifies images of 5 categories: fox, deer, wolf, brown bear and wild bear. Burghardt and Calic [18] created a real-time system which tracks the animals' head in the video and collects information about them. They applied the Viola-Jones detection algorithm which is basically utilized in human face detection.

Si et al. [19] presented a learning system for generative image representation which only needs a small number of training image samples. Each learned template consists of several image blocks and their properties like location, scale and orientation. These properties are represented by local features such as local sketch, texture gradients, flatness regions, and colours. Afterwards these blocks are ranked with respect to their information gain. Finally they performed automatic feature selection algorithm which makes the system adjustable for applying in different image classification applications. They evaluated the proposed system on different public benchmarks including LHI-Animal-Faces dataset.

## III. Backgrounds

Theory and essential components of DAG-CNN are presented in the following sub-sections. We briefly introduce the VGG-16 architecture which has been employed as the backbone of the proposed DAG-CNN approach.

### A. Directed Acyclic Graph-Convolutional Neural Network (DAG-CNN)

The directed acyclic graph (DAG) networks are capable of constructing more complicated network architectures compared to classic CNN which has a linear chain of layers [20]. DAG structure is inspired from the concept of recurrent neural networks (RNNs) which gets some feedback connections between forward and backward layers. These feedbacks will make the network to be able to capture the dynamic states. The major superiority of DAG architecture is its possibility of receiving multiple input parameters from several backward layers. As the result, it can gain different scales of image representation. The basic feature of deep neural networks (DNN), is the skip connections between layers that is alike the DAG-CNN's main idea and it is proven that these skip connections can improve the accuracy of the classification system accordingly.

Yang and Ramanan [21] have proposed the DAG-CNN in 2015. Their algorithm was applied to a collection of multi-scale images' features and tested for classification of 3 standard scene benchmarks. The authors proved that the multi-scale model can be carried out as a DAG-Structured feed forward CNN. According to this architecture, an end to end gradient-based learning can be applied for automatic multi-scale feature extraction by using generalized back propagation algorithm over the layers that have multiple inputs. Basically, the network training equations are following the standard CNN equations except for the ADD and RELU layers due to their several inputs and outputs.

Figure I presents the parameter setup for  $i^{th}$  ReLU layer, considering  $\alpha_i$  as its input and  $\beta_i^{(j)}$  as the output for its  $j^{th}$

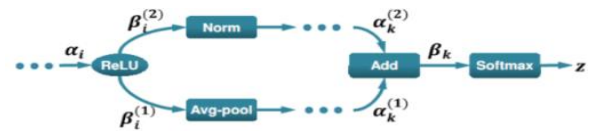


Figure I. Parameter configuration at  $i^{th}$  ReLU [21]

output branch (its  $j^{th}$  child in the DAG),  $z$  will be the final output of the softmax layer. The gradient of  $z$  regarding the input of the  $i^{th}$  ReLU layer can be calculated as follows:

$$\frac{\partial z}{\partial \alpha_i} = \sum_{j=1}^c \frac{\partial z}{\partial \beta_k} \frac{\partial \beta_k}{\partial \alpha_i} \quad (1)$$

where  $c$  indicates the number of output edge of the  $i^{th}$  ReLU.

Additionally,  $\beta_k = g(\alpha_k^{(1)}, \dots, \alpha_k^{(N)})$  represents the output of an Add layer with multiple inputs. By employing the chain rule, the gradient along the layer will be measured as stated in Eq. (2):

$$\frac{\partial z}{\partial \alpha_i} = \frac{\partial z}{\partial \beta_k} \frac{\partial \beta_k}{\partial \alpha_i} = \frac{\partial z}{\partial \beta_k} \sum_{j=1}^c \frac{\partial \beta_k}{\partial \alpha_k^{(j)}} \frac{\partial \alpha_k^{(j)}}{\partial \alpha_i} \quad (2)$$

In the convolutional layers, the convolution operation is computed by Eq. (3) as follows:

$$y(i, j) = \sum_{k=0}^{M-1} \sum_{l=0}^{N-1} x(i-k, j-l)h(k, l) \quad (3)$$

where  $x$  refers to input sample,  $h$  indicates the used filter, and  $M, N$  stand for the width and height of the input sample.



Figure IV. LHI-Animal-Faces dataset. Five images are shown for each category

The Convolution-layer's output is shown by  $y$ . In order to update neurons' biases and weights, Eq. (4)-(5) are applied to DAG-CNN layers:

$$\Delta W_l(t+1) = -\frac{x_l}{r} W_l - \frac{x}{n} \frac{\partial C}{\partial W_l} + m \Delta W_l(t) \quad (4)$$

$$\Delta B_l(t+1) = -\frac{x}{n} \frac{\partial C}{\partial B_l} + m \Delta B_l(t) \quad (5)$$

In the above equations,  $W, B, l, \lambda, x, n, m, t$ , and  $C$  signify the weight, bias, layer number, regularization parameter, learning rate, total number of training samples, momentum, updating step, and cost function, respectively.

In DAG-CNNs, considering the fact that the lower layers are directly linked to the output layer through multi-scale connections, it is assured that these layers' neurons receive a strong gradient signal during the learning process and will not suffer from vanishing gradients issue.

In CNNs, the dimension of the learned features in mid-layers can be very large. Therefore, concatenating these features may result in curse of dimensionality problem. To avoid this issue,

marginal activations are applied by operating the average pooling on the learned features of the layers which are used for score-level fusion.

### B. VGG-16 architecture

VGG-16 [22] which is proposed by the Oxford Visual Geometry Groups in ImageNet Large-Scale Visual Recognition Challenge (ILSVRC 2014) provides deeper and wider comparing to classic CNN architecture. VGG-16 consists of five batches of convolution operations; every batch can have 3 to 4 adjacent Conv-layers that are associated with max-pooling layers. The size of kernels in all convolutional layers is 3×3. Convolutional layers and the number of kernels have

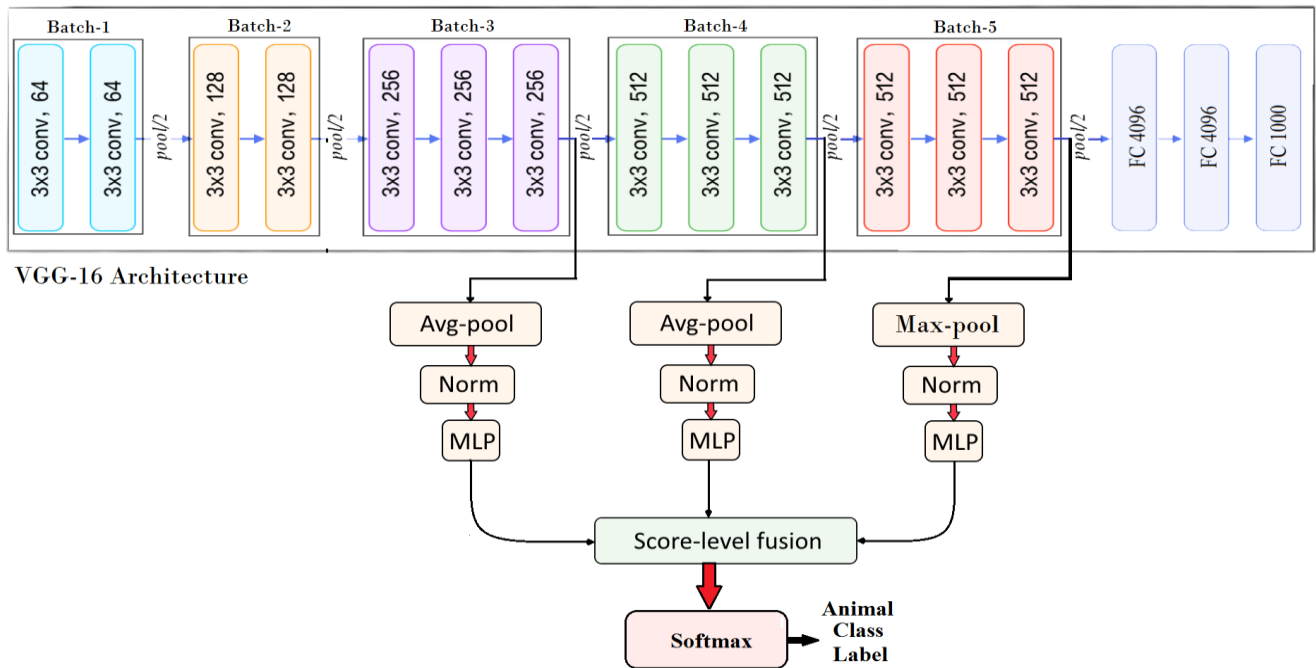


Figure III. Overview of proposed DAG-VGG16 method

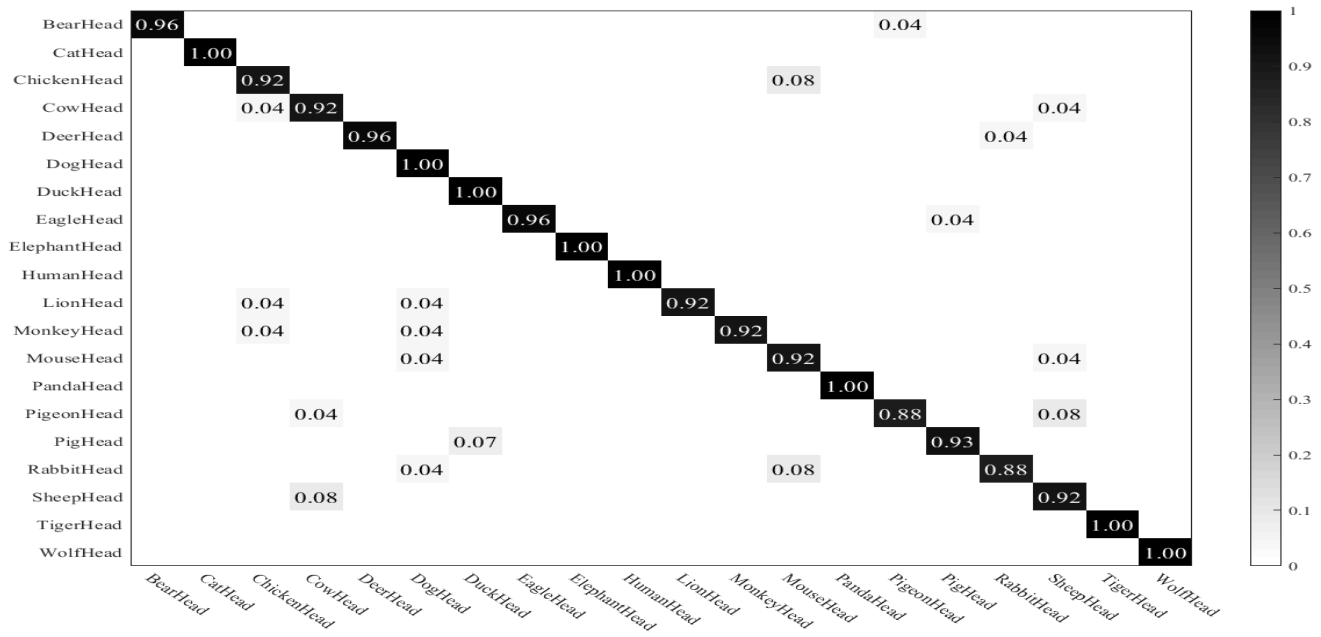
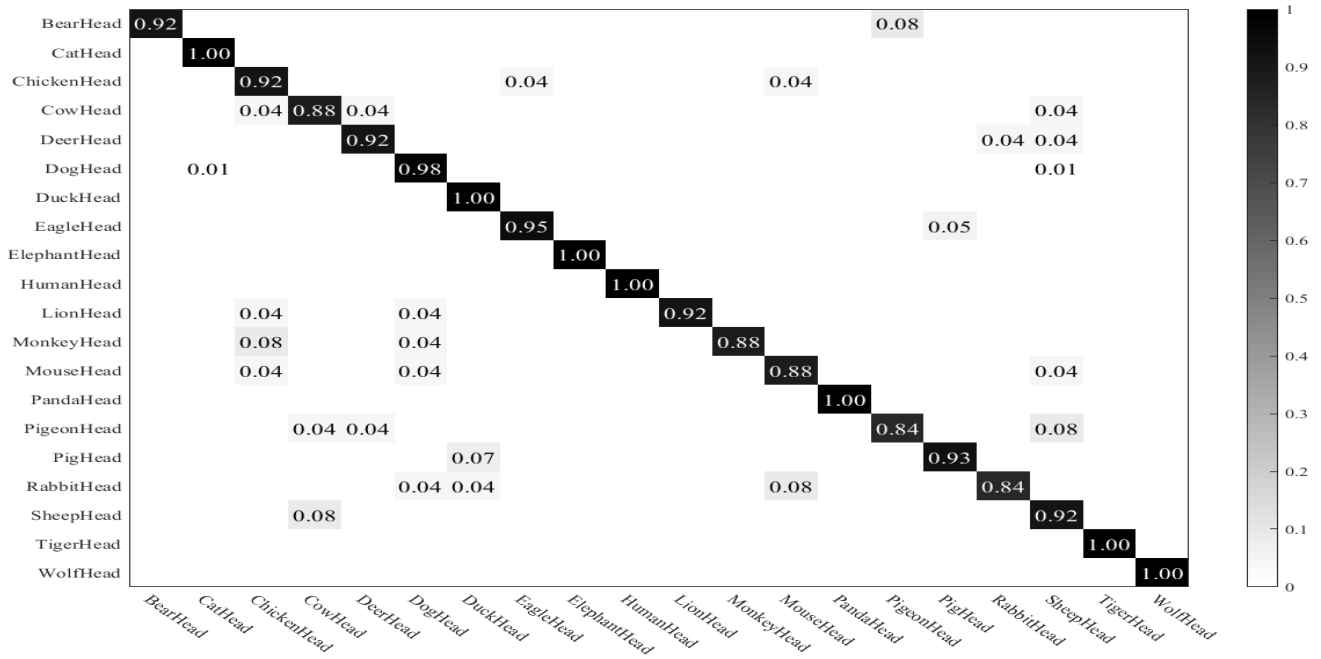


Figure V. Confusion Matrix. (up) Score-level fusion of VGG-16 and KFA[5], (down) DAG-VGG-16

similar sizes within each batch (starts from 64 in earlier layers and increases to 512 in the last one). Figure II demonstrates the 16-layer VGG architecture. This architecture has been employed in many research studies and in fact, it was the first architecture that outperformed the human-level performance on ImageNet [23].

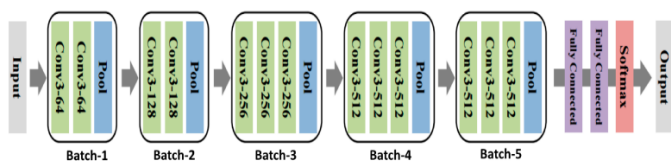


Figure II. VGG-16 Architecture

#### IV. Proposed Method

In this paper, a novel DAG-CNN architecture is proposed with the aim of animal classification from their head image samples by extracting the multi-stage features from multiple layers of the VGG-16 CNN model. Despite the fact that mid-level features of the CNN would be used to discriminate

patterns with different levels of details, in the most CNN-based approaches reported in the literature, these features are neglected.

The overview of the proposed method is illustrated in Figure III. In the pre-processing phase, some simple enhancement/operations such as image resizing, converting the RGB to grayscale and histogram equalization are done to remove the negative effect of various factors such as size,

illumination and image quality. Then two distinct set of features are generated. Afterwards, the similarity score between these feature vectors and the whole feature vectors available in the training set are computed and the minimum score is selected for each technique. The distance between the test and training animal images is considered as the score of that sample. At this point, the achieved scores are normalized and the fusion is performed by the addition of these

normalized scores and fed to Nearest Neighbour (NN) classifier to obtain the final animal's class label.

## V. Experiments and Results

The experiments employ LHI-Animal-Faces dataset [19] that contains 2200 head images of 19 different category of animals plus one class of human head. Five random samples from each of the aforementioned classes are shown in Figure IV. Due to large intra-class similarities and inter-class variations in the dataset, precise classification is a challenging task.

In order to achieve comparable result with state-of-the-art, we follow the experimental setup introduced by Si et al. [25]. On this setting, 30 random images from each class are selected for the training phase and the rest of images are utilized in the test stage.

The pre-processing step includes image resizing to  $224 \times 224$  pixels, pixel intensity normalization and histogram equalization. VGG-16 is selected as the underlying architecture of the proposed DAG-CNN due to its success in

Table I. Classification Accuracy on LHI-Animal-Faces dataset

Method	Accuracy
HOG+SVM[19]	70.8%
SW-RBF[25]	44%
FRAME[26]	79.4%
HIT[19]	75.6%
LSVM[27]	77.6%
AOT[24]	79.1%
Deep Boosting[13]	81.5%
Score-level fusion[7]	95.31
Proposed DAG-CNN	96.29%

image classification. By following the greedy approach explained in [2], three new links are added to Batch-3, Batch-4 and Batch-5 (Figure III). The proposed system is pre-trained by hundreds of thousands images from large public image repository, namely ImageNet database [28] and learned powerful discriminative feature sets.

In the next step by utilizing the target dataset, we fine-tuned the pre-trained proposed system. Data augmentation techniques like translation, rotation and flipping are applied to avoid overfitting and the training size is expanded five times. The first early layers' weights are frozen so that they remain

## References

- [1] Elson, J., Douceur, J., Howell, J., et al.: Asirra: a CAPTCHA that exploits interest-aligned manual image categorization. In: Proc. of ACM CCS, pp. 366-374. (2007)
- [2] Taheri, S., Toygar, Ö.: On the use of DAG-CNN architecture for age estimation with multi-stage features fusion. *Neurocomputing*. 329,300-310 (2018)

intact throughout the fine-tuning process. The remaining layers' weights are fine-tuned by applying Stochastic Gradient Descent (SGD) algorithm to minimize the loss function with a small initial learning rate of 0.001.

In order to investigate the efficiency of the proposed DAG-CNN system, we compare its performance with the-state-of-the-art algorithms that utilized the LHI-Animal-Faces dataset. This comparison is summarized in Table 1 and shows that the proposed method outperforms other state-of-the-art algorithms for animal face classification. In [7], we examined several categories of feature extractors including hand-crafted local and appearance-based descriptors and automatic learned features. We applied well-known local feature descriptors such as Histogram of Oriented Gradients (HOG), Completed Local Binary Patterns (CLBP), LBP histogram Fourier (LBP-HF), Haralick features and Median Robust Extended LBP (MRELBP). In order to utilize the discrimination power of appearance-based descriptors, Linear Discriminant Analysis (LDA) and Kernel Fisher Analysis (KFA) are examined. Finally we tested the learned features of pre-trained and fine-tuned publicly available CNN architecture, namely, AlexNet and VGG-16. We investigated score-level fusion of various combination of the aforementioned descriptors and we showed that fusion of hand-crafted features and multi-stage CNN-based features gain even higher accuracy comparing to CNN. On the other hand, the proposed DAG-CNN system causes meaningful improvement in accuracy. The 96.4% classification accuracy was achieved by automatically combining multi-stage learned features. This result is superior to all of the state-of-the-art methods mentioned in Table 1, even our previous system which utilized hand-crafted and learned feature fusion.

The confusion matrix is computed to assess the classification precision for the two best results of Table 1 and is shown in Figure V. For 8 animal categories, the proposed system successfully classified the images with 100% accuracy and for the 10 classes, the classification precisions are higher than 92%. The maximum confusions are generated by the sheep category versus cow head, rabbit and chicken head versus mouse head (8%).

## VI. Conclusions

In this paper, a new architecture of a nonlinear CNN structure, namely DAG-CNN is presented for animal classification task. This system automatically combines different layers' features of a CNN for making decision. For this purpose, several new links are added to the underlying backbone from a popular CNN structure, namely VGG-16. The proposed method is compared with several state-of-the-art systems that use the same dataset. The comparison results show that the proposed DAG-CNN architecture outperforms the state-of-the-art systems for animal face classification.

- [3] Druzhkov, P. N., Kustikova, V. D.: A survey of deep learning methods and software tools for image classification and object detection. *Pattern Recognition and Image Analysis*. 26(1), 9–15 (2016)
- [4] Gao, H., Cheng, B., Wang, J., Li, K., Zhao, J. and Li, D.: Object Classification Using CNN-Based Fusion of Vision and LIDAR in Autonomous Vehicle Environment. *IEEE Transactions on Industrial Informatics*. 14(9), 4224-4231 (2018).

- [5] Golrizkhatami, Z., Acan, A.: ECG classification using three-level fusion of different feature descriptors. *Expert Systems with Applications*. Vol 114, pp.54-64(2018)
- [6] Taheri, S., Toygar, Ö.: Multi-stage age estimation using two level fusions of handcrafted and learned features on facial images. *IET Biometrics*, 8(2) , pp. 124 – 133 (2019)
- [7] Taheri, S., Toygar, Ö.: Animal classification using facial images with score-level fusion. *IET Computer Vision*. 12(5), 679 – 685 (2018)
- [8] Schmid, C.: Constructing models for content-based image retrieval. In: *CVPR '01*, Kauai, United State. pp. 11-39. *CVPR* (2001)
- [9] Ramanan, D., Forsyth, D. A., Barnard, K.: Detecting, localizing and recovering kinematics of textured animals. In: *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, June 2005, San Diego, USA, pp. 635-642. *IEEE* (2005)
- [10] Ramanan, D., Forsyth, D.A., Barnard, M.-K.: Building models of animals from video. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28 (8), pp. 1319-1334. *IEEE* (2006)
- [11] Berg, T. L., Forsyth, D. A.: Animals on the web. *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06)*, NY, USA, pp. 1463-1470. *CPVR* (2006)
- [12] Cao, Z., Principe, J.C., Ouyang, B., Dalgleish, F., Vuorenkoski, A.: Marine animal classification using combined CNN and hand-designed image features. In *OCEANS15 MTS/IEEE Washington*, pp. 1-6. *IEEE* (2015)
- [13] Penga, Z., Lia, Y., Caib, Z., et al.: Deep Boosting: Joint feature selection and analysis dictionary learning in hierarchy. *Neurocomputing*. 178(20), 36–45 (2016)
- [14] Afkham, h., Tavakoli, A., Eklundh, J., Pronobis, A.: Joint Visual Vocabulary For Animal Classification. In: *ICPR 2008*, Tampa, FL, USA, pp. 1-4. *ICPR* (2008)
- [15] Kumar, Y.S., Manohar, N., Chethan, H.K.: Animal classification system: a block based approach. *Procedia Computer Science*, 45, 336-343 (2015)
- [16] Manohar, N., Kumar, Y.S., Kumar, G.H.: Supervised and unsupervised learning in animal classification. In: *Advances in Computing, Communications and Informatics (ICACCI)*, International Conference on, pp. 156-161. *IEEE*(2016)
- [17] Matuska, S., Hudec, R., Benco, M., Kamencay, P., Zachariasova, M.: A novel system for automatic detection and classification of animal. In: *ELEKTRO*, pp. 76-80. *IEEE* (2014)
- [18] Burghardt, T., Calic, J.: Real-time face detection and tracking of animals. In: *Neural Network Applications in Electrical Engineering. NEUREL 2006. 8th Seminar on*, pp. 27-32. *IEEE*(2006)
- [19] Si, Z., Zhu, S.-C.: Learning hybrid image templates (hit) by information projection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 34(7), 1354–1367 (2012)
- [20] Golrizkhatami, Z., Taheri, S., Acan, A.: Multi-scale features for heartbeat classification using directed acyclic graph CNN. *Applied Artificial Intelligence* 32 (7-8), pp. 613-628 (2018)
- [21] Yang, S., Ramanan, D.: Multi-scale recognition with DAG-CNNs. In: *Computer Vision (ICCV), 2015 IEEE International Conference*, pp. 1215-1223. *ICCV* (2015)
- [22] Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556* (2014)
- [23] Szegedy, C., Liu, W., Jia, Y., et al.: Going deeper with convolutions. *CVPR* (2015)
- [24] Si, Z., Zhu, S.-C.: Learning and-or templates for object recognition and detection, *Pattern Analysis and Machine Intelligence*. *IEEE Transactions on*. 35 (9), 2189–2205 (2013)
- [25] Kolouri, S., Zou, Y., Rohde, G.K.: Sliced Wasserstein kernels for probability distributions. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 5258-5267. *IEEE* (2016)
- [26] Xie, J.X.: *Generative Modeling and Unsupervised Learning in Computer Vision*. Doctoral dissertation, *UCLA University*, (2016)
- [27] Felzenszwalb, P., Girshick, R., McAllester, D., Ramanan, D.: Object Detection with Discriminatively Trained Part-Based Models. *IEEE Trans. Pattern Analysis and Machine Intelligence*. 32(9), pp.1627-1645 (2010)
- [28] Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K. Fei-Fei, L.: Imagenet: A large-scale hierarchical image database. In *Proc. CVPR*, 2009.

# Movie Success Prediction with Statistical Analysis Techniques and Machine Learning Methods

Bugay Sarıkaya  
Department of Computer  
Engineering  
Baskent University  
Ankara, Turkey

Duygu Dede Sener  
Department of Computer  
Engineering  
Baskent University  
Ankara, Turkey

**Abstract**— In the movie industry, huge investments have been made to shoot a successful motion picture. However, despite large investments, there are some movie examples that cannot be successful as expected. Therefore, predicting the success of a movie is so important on a large scale for the movie producers before releasing the movie. In this study, a classification-based prediction model is aimed to develop for providing a foresight to the producers about investing on a movie. Different statistical analysis and machine learning approaches were used in the proposed model for predicting success of a movie. We mainly focus on detecting which movie attribute is highly correlated with the success of the movie and which machine learning technique is better at predicting the movie success. To do so, firstly a statistical analysis was conducted by using chi-square analysis and analysis of variance test. Then a comparative analysis was performed by using different machine learning techniques including random forest, support vector machine and artificial neural network. The experimental results indicate that the most important predictors of a movie's success are "voteAverage", "voteCount", "revenue" and "budget. In addition to this, random forest has become successful by the accuracy of 96% in predicting movie success among other machine learning methods.

**Keywords**— *Movie success prediction, Movie success classification, Machine Learning, Model Prediction*

## I. Introduction

Movie industry is one of the expanded industries all around the world. With the rapid growth of this industry and economic impact of it, many researchers have been studying on the movie industry. Especially conducting predictive models to investigate factors which has impact on the success of a movie has become a popular research area over the past decade. It is a well-known fact that the success of a motion picture is based on several features of the movie such as box-office, budget, revenue, and popularity level. Although, there are some other important factors such as director, actors which has undeniable impact on a movie's success, every movie cannot achieve the expected box-office or the success every time. Therefore, there is a basic need for producers to predict the success of a movie before its release. Recently there are some studies which focus on the predicting the success of movie using different machine learning approaches.

Ahmad et al. [1] proposed a mathematical model for predicting movie success including finding correlations between various features using X2 analysis. Simulation data was used in the study and only tested in Bollywood movies and they also concluded that actors and film genres affect the success of the film. Ping-Yu Hsu et al. [2] developed a special model to predict user ratings with IMDB (Internet Movie Database) attributes. There are 32968 films in the used dataset

and linear combination, multiple linear regression and neural network methods and X2 analysis were used. It is stated that some attributes, such as writers, actors, and directors, profoundly affect user ratings. In addition to this, Eker et al. [3] sought to measure the effect of features on movie classification. Decision trees, K-NN, Random Forest, c4.5, c5.0 and Boosting algorithms were used in this study. A comparative study was performed by using different machine learning algorithms on the dataset from IMDB website and Facebook website. It has been observed that the user votes are the most important factor in the IMDB score, and the country where the film is produced is the least important factor in determining the IMDB score. Saraee et al. [4] studied on IMDB data using various data mining techniques. Furthermore, Lash and Zhao [5] proposed a way to predict decisions regarding film investments. This study assisted in making an early investment decision in filmmaking by using a historical data. In this study, the profit was calculated mainly on the box-office revenue. However, for many movies, there are other sources of income, such as items for sale. Kyuhan Lee et al. [6] examine multiple approaches to improve the performance of the prediction model. They developed and added a new feature derived from the theory of transmedia storytelling. They used an ensemble approach, which has rarely been adopted in the research on predicting box-office performance. As a result, the proposed model, Cinema Ensemble Model (CEM), outperformed the prediction models from the past studies that use existing machine learning algorithms. Besides this studies, Hemraj Verma and Garima Verma [7] conducted a comparative analysis of prediction models using various machine learning techniques. The models were used to predict whether a movie would be a hit or a flop before it came out. The major predictors used in the models are the ratings of the lead actor, IMDB ranking of a movie, music rank of the movie, and the total number of screens planned for the release of a movie.

In this study, statistical analysis techniques and different machine learning methods were performed to predict the success of a movie. Two statistical tests such as chi-square analysis and analysis of variance test (ANOVA) and three different machine learning algorithms including random forest (RF), support vector machine (SVM) and artificial neural network (ANN) were performed on a collected dataset. A comparative study was performed to investigate which algorithm achieve the highest rate of the accuracy. An experimental study is performed on a dataset collected from IMDB dataset. According to the results the most important predictors of a movie's success are "voteAverage", "voteCount", "revenue" and "budget. Besides this, it has been found that random forest is the most successful technique by the accuracy

of 0.96 in predicting movie success among other machine learning methods.

## II. Methods

In this section used statistical analysis techniques and machine learning techniques in the proposed model are described.

### A. Statistical Analysis Methods

#### a) Correlation Matrix

Attribute correlation matrix was used to obtain correlation between each attribute and the target attribute which is movie success in our study. It is a common way to summarize data and guide the data owner to focus more on highly correlated attributes in which way data analysis can be conducted in more efficient way. In our study, Pearson correlation coefficient was used. The score ranges between 0 and 1, the values close to 1 represents a high correlation while the values close to 0 represents a low correlation.

#### b) Chi-Squared Test

A chi-square ( $\chi^2$ ) test [8] is a hypothesis testing method. It is used for comparing the observed values with the expected values to detect the stated null hypothesis is true or not. The null hypothesis states that there is no difference between the compared data. For this test, a p-value that is less than or equal to the defined significance level (0.05) indicates there is a strong evidence to conclude that the observed distribution is not the same as the expected distribution. Moreover, the data used in calculating a chi-square statistic must be random, raw, mutually exclusive, drawn from independent variables, and drawn from a large enough sample.

#### c) Analysis of Variance (ANOVA)

Analysis of variance (ANOVA) [9] is an analysis tool used in statistics that splits an observed aggregate variability found inside a data set into two parts: systematic factors and random factors. The systematic factors have a statistical influence on the given data set, while the random factors do not. Analysts use the ANOVA test to determine the influence that independent variables have on the dependent variable in a regression study. In our study, two-way ANOVA was used because we wanted make comparisons between the means of three groups of data, where two independent variables are considered. The considered variables are movie success and the rest of the attributes. Moreover, Multivariate ANOVA (MANOVA) was used to extend the capabilities of analysis ANOVA by assessing multiple dependent variables simultaneously.

### B. Machine Learning Methods

#### a) Random Forest

Random forests (RF) or random decision trees [10] are combination of tree predictors such that each tree depends on the values of a random vector sampled independently and with the same distribution for all trees in the forest. It is an ensemble learning method used for classification, regression and other tasks which needs constructing a multitude of decision trees. For the classification task, each generated tree predicts a class as an output class and the final decision is

made by selected mostly predicted classes from the trees. RFs and its variants are called as black-box models and they have been applied on a variety of research fields such as bioinformatics, finance and healthcare systems.

#### b) Support Vector Machine (SVM)

Support vector machine (SVM) [11] is a supervised machine learning algorithm that used for data classification and regression analysis. SVM's main goal is finding a hyperplane that best divides a dataset to two different classes multiple times (as many times needed to match number of classes). Support vectors are the data points nearest to the hyperplane, this point help define the hyperplane, so all computations are done through these points. This hyperplane creates an area of margin which divides two classes apart. Error function here is designed so that the margin becomes larger as error decreases. If there is no clearly dividing hyperplane then the whole feature space is transformed into a new higher dimension feature space. This is known as kernelling. SVMs produce accurate results on clean datasets with small to medium sample size. When dealing with larger datasets however computational costs can be too much to handle and also it is highly sensitive to the noisy nature of large datasets.

#### c) Artificial Neural network (ANN)

An artificial neural network (ANN) [12] is the piece of a computing system designed to simulate the way the human brain analyzes and processes information. It is the foundation of artificial intelligence (AI) and solves problems that would prove impossible or difficult by human or statistical standards. ANNs have self-learning capabilities that enable them to produce better results as more data becomes available. Artificial neural networks are built like the human brain, with neuron nodes interconnected like a web. The human brain has hundreds of billions of cells called neurons. Each neuron is made up of a cell body that is responsible for processing information by carrying information towards (inputs) and away (outputs) from the brain.

## III. Results

In this section, the used dataset is explained and experimental results are given in the following sections.

### A. Dataset

In this study, a dataset of 4899 movies released from 2000 to 2020 was collected from TMDB (The Movie Database) [13] and OMDB (The Open Movie Database) [14]. The dataset consists of various type of attributes such as date, genre, language, season, IMDB (Internet Movie Database) rating of the features are categorical, box-office, budget, IMDB votes, popularity, revenue, run time of them are numerical values. In addition to this, the sample distribution of the dataset there are 1215 samples for successful class, 2663 samples for average successful and 1021 samples for unsuccessful class.

#### a) Data Preprocessing Steps

In the dataset, each movie has a IMDB rating for representing the success of the movie. Unlike a binary classification problem, having two classes, multi-class

classification would cover the movie success information more broadly. An artificial class construction has been performed to transform the problem into a multi-class classification problem. Therefore, each movie is categorized based on its IMDB rating such as movies that have the scores in the -range of [7-10] were assigned to class “successful”, the scores in the range of [5-6.99] assigned to the “average successful” class, then the scores in the range of [0-4.99] was matched with the “unsuccessful” class. In this way, our problem has been converted into a multi-class problem.

Furthermore, missing value removal and transforming categorical features into numeric features were performed before applying classification algorithms. One-hot encoding was applied to convert categorical values into the numerical values. In this way, a generalized form to be provide to machine learning algorithms can be obtained. Finally, feature scaling approach was applied on some attributes to normalize the range of independent attributes of data like box-office, revenue and vote counts.

### B. Evaluation Metrics

To evaluate the performance of the classification algorithms one metric was used: accuracy. Accuracy (1) refers to the degree of conformity of a measured or calculated quantity to an actual (true) value. A result is said to be accurate when it matches to a particular target. TP, FP, TN and FN values represent true positives, false positives, true negatives and false negatives, respectively. Overall accuracy were used to compare the performance of classification algorithms.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

### C. Experimental Results

To get insight about the success distribution among the movies based on the attributes, some distribution plots were obtained. Figure 1 gives the movie success distribution based on the genre attribute. The class categories are given with different colors such as red, green and blue represent unsuccessful, average and successful movie class respectively. According to the figure the highest number of successful movies can be seen in “drama” and “comedy” movies, while most of the “horror” and” thriller” movies are in the unsuccessful class.

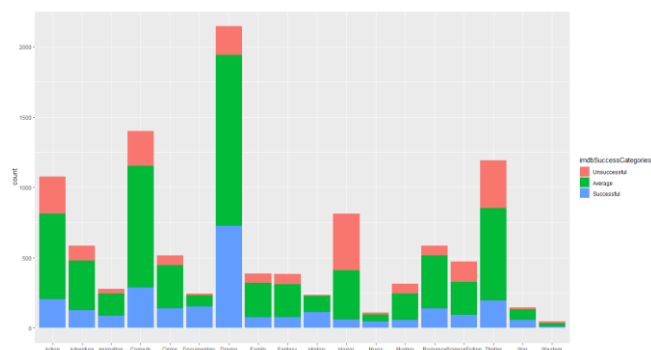


Figure I. Chart of genre by success of movie

Then correlation matrix was obtained given in Figure 2. According to the matrix, the movie success given as the name “imdbRating” is highly correlate with the attributes which are “voteAverage”, “voteCount”, “revenue” and “budget”. After having obtained highly correlated attributes with the success statistical analysis has been performed by focusing on these obtained attributes to conclude the relationship between them. To do so, chi-square ( $\chi^2$ ) test, ANOVA and MANOVA tests were performed to investigate which movie attribute has the biggest impact on the movie success.

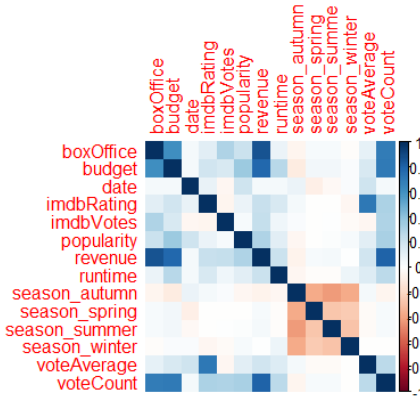


Figure II. Correlation matrix of movie attributes

Chi-square ( $\chi^2$ ) test statistic and corresponding p-values are given in Table 1. According to the obtained p-values which are all less than the significance level of 0.05, we can conclude that the association between each attribute and movie’s success is statistically significant but excluding the attribute of “season\_summer”, since it does not have the p-value less than or equal to 0.05. In addition to this, “Df” value in the table represents “degrees of freedom”.

Table I. Chi-Squared test result

Attribute	X-squared	Df	p-value
voteAverage	9645.6	156	< 2.2e-16
voteCount	8865.4	3222	< 2.2e-16
revenue	10081	4310	< 2.2e-16
budget	5728	2154	< 2.2e-16
runtime	3230	372	< 2.2e-16
boxOffice	9119.6	3978	< 2.2e-16
date	125.36	20	< 2.2e-16
popularity	19113	8300	< 2.2e-16
season_winter	14.283	2	0.0007914
season_autumn	14.283	2	0.0007914
season_summer	4.9467	2	0.0843
season_spring	17.391	2	0.0001674
imdbVotes	4825.6	1512	< 2.2e-16

Two-way ANOVA results are given in the Table 2. The attributes in table are selected based on their correlation score and chi-square statistic results. According to the results the difference between the means of group, which are successful, average successful and unsuccessful movies, are statistically significant, because all obtained p-values are less than the significance level of 0.05. It can be concluded that all selected attributes. are important factors to emphasize the success of a movie.



Table II. Two-way ANOVA results

Two Way with IMDB Success Categories	Df	Sum Sq	Mean Sq	F Value	Pr(>F)
voteAverage	2	7449	3725	4062	<2e-16
voteCount	2	8.003e+09	4.001e+09	592.8	<2e-16
revenue	2	1.532e+19	7.662e+18	259.1	<2e-16
budget	2	6.593e+17	3.296e+17	192.1	<2e-16
runtime	2	224351	112175	163.1	<2e-16
boxOffice	2	9.171e+17	4.585e+17	129.5	<2e-16

Besides the ANOVA results, MANOVA was also conducted to assess multiple dependent variables simultaneously. “voteAverage” and “voteCount” attributes are selected attributes for this task. According to the results it can be concluded that difference between the means of group has been found statistically significant with grouping these attributes.

Table III. Multivariate ANOVA (MANOVA) results

MANOVA with IMDB Success Categories	Df	Pillai	approx F	num Df	den Df	Pr(>F)
voteAverage / voteCount	2	0.45203	1596.1	4	21864	< 2.2e-16

After having had some statistical analysis observations about the dataset, the movie success prediction was performed by using three different machine learning algorithms. Accuracy metric was used to compare performance of the used methods. Moreover, 10-fold cross validation technique was conducted to obtain a more accurate model by providing each sample or movie to take part in the training and testing of the proposed model. One-versus-all method was used in order to extend binary-class problem into multi-class problem. All experiments were performed by using R language on RStudio.

To perform ANN method a layered structure was prepared as given in Figure 3. The black lines show the connections between each layer and the weights on each connection while the blue lines show the bias term added in each step. The bias can be thought as the intercept of a linear model. Table 3 gives the average accuracy values of all three classes. According to table, Random Forest method has become most successful method in predicting movie success with the accuracy rate of 0.96. It can be also observed that the lowest accuracy belongs to SVM method.

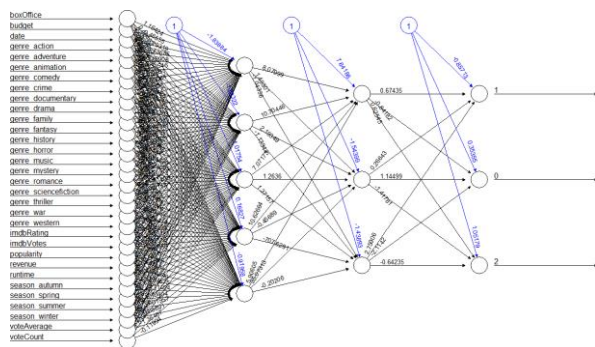


Figure III. Graphical representation of the used ANN model with the weights on each connection

Table IV. Classification performances of the machine learning algorithms

	Method		
	RF	SVM	ANN
Accuracy	0.96	0.89	0.94

#### IV. Conclusion

With the rapid growth of the movie industry, large investments have been made to release successful movies. It is also a well-known fact that the movie success does not only depend on one or two criteria, but so comprehensive studies should also be conducted to investigate all criteria which is related with the movie success. Therefore, knowing these success criteria before the movie is released, it is very important for the filmmakers to shoot the movie.

In this study, we present a model for predicting the success of a movie by using different statistical analysis techniques and machine learning approaches. We mainly focus on detecting which feature of a movie is highly correlate with the success of the movie and which machine learning technique is better at predicting the movie success. According to the experimental results most important predictors of a movie’s success are “voteAverage”, “voteCount”, “revenue” and “budget”. In addition to this, random forest has become successful by the accuracy of machine learning algorithms in predicting movie success among other machine learning methods.

The proposed study as a supplementary tool can provide a practical implication for the decision makers and producers in the movie industry. They can make a decision about investment on a movie before its releasing. In the future, the current dataset can be expanded by adding new movies and attributes to get more reliable and robust predictive model.

#### References

- [1] Ahmad, J., Duraisamy, P., Yousef, A., & Buckles, B. (2017, July). Movie success prediction using data mining. In 2017 8th International Conference on Computing, Communication and Networking Technologies (ICCCNT) (pp. 1-4). IEEE.
- [2] Hsu, P. Y., Shen, Y. H., & Xie, X. A. (2014, October). Predicting movies user ratings with imdb attributes. In International Conference on Rough Sets and Knowledge Technology (pp. 444-453). Springer, Cham.
- [3] Eker, A. G., Duru, N., Kat, O., & Ildırar, A. (2018, September). Makine Öğrenmesi ile Film Başarı Tahmini. In 2018 3rd International Conference on Computer Science and Engineering (UBMK) (pp. 610-614). IEEE.
- [4] Saraee, M. H., White, S., & Eccleston, J. (2004). A data mining approach to analysis and prediction of movie ratings. Transactions of the Wessex Institute, 343-352.
- [5] Lash, M. T., & Zhao, K. (2016). Early predictions of movie success: The who, what, and when of profitability. Journal of Management Information Systems, 33(3), 874-903.
- [6] Lee, K., Park, J., Kim, I., & Choi, Y. (2018). Predicting movie success with machine learning techniques: ways to improve accuracy. Information Systems Frontiers, 20(3), 577-588.
- [7] Verma, H., & Verma, G. (2020). Prediction model for bollywood movie success: A comparative analysis of performance of supervised machine learning algorithms. The Review of Socionetwork Strategies, 14(1), 1-17.

- [8] Plackett, R. L. (1983). Karl Pearson and the chi-squared test. *International Statistical Review/Revue Internationale de Statistique*, 59-72.
- [9] Scheffe, H. (1999). *The analysis of variance* (Vol. 72). John Wiley & Sons.
- [10] Breiman, L. (2001). Random forests. *Machine learning*, 45(1), 5-32.
- [11] Noble, W. S. (2006). What is a support vector machine?. *Nature biotechnology*, 24(12), 1565-1567.
- [12] Wang, S. C. (2003). Artificial neural network. In *Interdisciplinary computing in java programming* (pp. 81-100). Springer, Boston, MA.
- [13] <https://www.themoviedb.org/> access date: 20.04.2021
- [14] <https://www.omdbapi.com/> access date: 20.04.2021

# Digital Controllers Performance Analysis for a Robot arm

Abdullah Ahmed Al-dulaimi  
Department of Electrical Electronics  
Engineering  
Karabuk University  
Abdalluhahmed1993@gmail.com

Mohammed Majid Abdulrazzaq  
Department of Computer Engineering  
karabuk University  
moh.abdulrazzaq9@gmail.com

Mustafa Mohammed Alhassow  
Department of Electrical and Computer  
Engineering  
Altinbas University  
Mustafa.alshakhe@gmail.com

Noor Qasim AL saedi  
Department of Electrical Electronics  
Engineering  
Karabuk University  
Noorqasimat@gmail.com

**Abstract**—The design methodology and performance study of various forms of digital compensators for a robot arm joint control system with sensor input are presented in this article. continuous time (s-plane or w-plane) and Discrete (z-plane) domain parameters are used in the design process. The frequency response characteristics design techniques were investigated, and five basic types of controllers were modelled and simulated using MATLAB: phase-lag, phase-lead, proportional-integral (PI), proportional-derivative (PD), and proportional-integral derivative (PID). Many of the controllers have been set up to maintain a 40-degree step margin. both closed loop phase answers as well as open loop bode plots have been analyzed. This paper presents a comparison of the controllers based on their phase reaction characteristics.

**Keywords:** digital controllers, PID controllers, robot arm, robot arm controllers, Digital Controllers Performance.

## I. Introduction

Controllers are needed to assess adjustments in system parameters and to meet performance requirements for steady-state precision, transient reaction, reliability, and disturbance prevention. Analog control systems are stable, with no intrinsic band width limitations or system changes. Due to the tolerances of practical machines, In analog controls, intricate logics are difficult to synthesize, while rendering complex interfaces among multiple subsystems is very difficult, and are vulnerable to incorrect designs and limitations. Furthermore, extraneous noise sources can corrupt analog systems significantly. High-tech digital controls Since no signal loss occurs in an along to digital (A/D) and digital to analog (D/A) conversions [1], systems are reliable. Furthermore, with a more sophisticated logic implementation, systems are more flexible and accurate. Filters do not encounter external noises, which makes them well-suited for adaptive filtering uses. Fast response and a digital memory interface are possible for digital systems [1]. A physical planet or system is accurately controlled through closed-loop or feedback operation where an output (system response) is adjusted as required by an error signal [2]. The discrepancy between the sponge as determined by sensor input and the target response generates the error signal. The error quantity is processed by a controller or compensator in order to satisfy those output requirements [3]. This paper describes five digital controller design methodologies for a robot control system in real time. In these design methods, the compensating parameter is the phase margin specified in the plant bode diagram. The design method employs strategies of frequency response that allow for frequency cross-over phase margin (Pm). Phase-lag, PI and PID controllers (lag, lead) were drawn up in accordance with the principle of compensation and

the methodologies as defined in [4]. This paper clarified the statistical and conceptual pricing articulated in the references. The basic and illustrative frames and approaches to digital control systems are mentioned in [6]. Digital control systems have been documented for training, theory, simulation and experimental approaches [7],[8]. A closed loop model has been introduced in [1] and [4] for digital systems control and implementations in the Digital Drive Controller. Regarding to PID controller we citation the information from [5].

## II. Methodology

A sampler, D/A block that is a zero-order hold (ZOH), a servomotor represented by an s-domain transfer mechanism, digital controller block, a power amplifier gain, gears represented by a gain value, and a feedback sensor block comprise the example robot control scheme outlined in this article. A s-domain transfer function presents the uncompensated plant. A/D conversion is started by the sampler and D/A conversion is held at zero order. Controllers must offset the margin of the plant phase and the desired result shall be 40 deg. For each controller, steady-state error, percent overshoot, rise time, and settling time are calculated for output assessment. This paper section by section documents a literature review of digital compensation, an example uncompensated robot arm joint plant, discrete and continuous time equations with design method, MATLAB simulation results of lag, lead, PI, PD and PID controls, and a comparative study among these five. In order to evaluate design requirements, the digital system was adopted, simulated and extended in MATLAB.

## III. Literature Review

The compensation theory, plant configuration and the mathematical derivations of design approaches, loop parameters and open loop of the controllers mentioned in this paper fully follow the literature provided in [1]. The controller transfer function for first-order compensation can be written as

$$D(z) = \frac{K_d(z-z_0)}{z-z_p} \quad (1)$$

Here,  $z_0$  and  $z_p$  represent the zero and pole positions, respectively. The controller's bilinear or trapezoidal transformation from the discrete z-plane to the continuous w-plane (warped s-plane) implies  $D(w) = D(z), z = \frac{1+(T/2)w}{1-(T/2)w}$  and  $D(w) = a_0 \frac{1+(w/\omega_{w0})}{1+(w/\omega_{wp})}$ ,

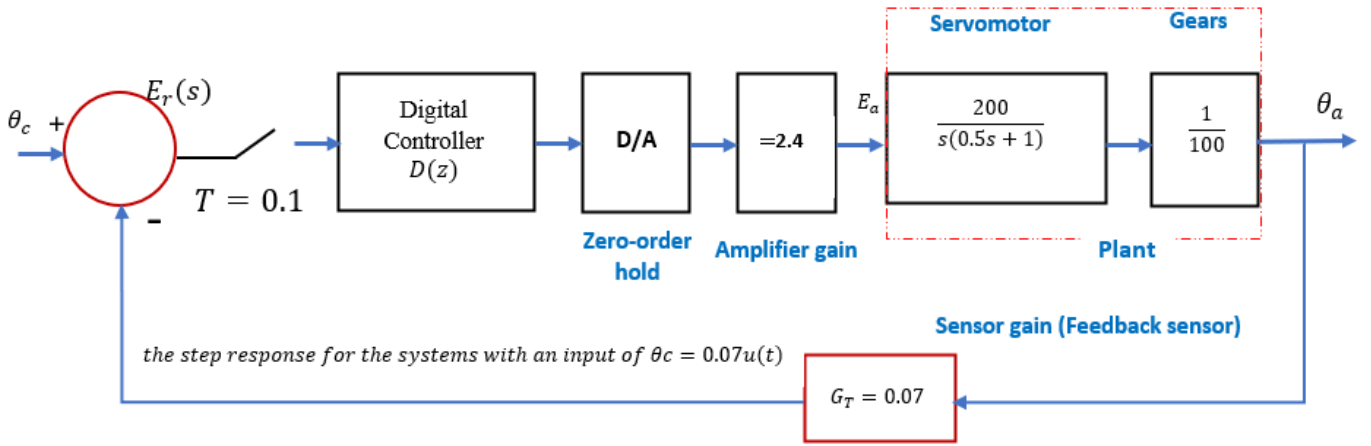


Figure 1. Robot arm joint control system block diagram

Here  $\omega_{w0}$  and  $\omega_{wp}$  denotes the zero and pole positions in the w-plane, and  $a_0$  denotes the compensator dc gain. The bilinear approximation states that:

$$w = \frac{2z-1}{Tz+1} \quad (2)$$

From the equations (1)-(4), in z-plane the controller can be realized as

$$D(z) = a_0 \frac{\omega_{wp}(\omega_{w0}+2/T) z - \frac{2/T-\omega_{w0}}{2/T+\omega_{w0}}}{\omega_{w0}(\omega_{wp}+2/T) z - \frac{2/T-\omega_{wp}}{2/T+\omega_{wp}}} \quad (3)$$

The equation (1) yields to

$$K_d = a_0 \frac{\omega_{wp}(\omega_{w0}+2/T)}{\omega_{w0}(\omega_{wp}+2/T)} \quad (4)$$

$$z_0 = \frac{2/T-\omega_{w0}}{2/T+\omega_{w0}} \quad (5)$$

$$z_p = \frac{2/T-\omega_{wp}}{2/T+\omega_{wp}} \quad (6)$$

#### IV. block diagram explanation

A closed loop model for digital control systems and applications of digital controllers to speed drives has been shown in the above diagram. Thus, consists of a sampler, digital controller block, D/A block which is a zero-order hold (ZOH), a power amplifier gain, a servomotor represented by a s-domain transfer function, gears represented by a gain value and a feedback sensor block. In the case of a closed-loop feedback system, the  $D(z)$  digital controller system is implemented. The controller uses algebraic algorithms such as filters and compensatory controls to correct or regulate the controlled system's behavior. The zero-order hold is a practical mathematical model of signal reconstruction using a digital-to-analog converter (ZOH). This can be illustrated by you take a and convert it to a continuous-time signal, at a set time, it stores each sample value and doesn't allow changes. The amplitude or power of a signal input to output port can be increased by connecting it to an amplifier whose gain is set to a particular level [9]. In order for a servomotor to transform the control signal from the controller into the rotational angular displacement or angular velocity of the motor output shaft, implies it has a servomotor. To power the arms of the robot,

Servo motors are used. Gears are used to transfer motion. By finding the torque and speed of the output gear, you can find the torque and speed of the input gear. The uncompensated plant is presented by a s-domain transfer function. The sampler initiates A/D conversion and zero-order hold implements D/A conversion. For performance evaluation, steady-state error, percent overshoot, rise time and settling time are measured for each controller. Here the given sensor feedback gain,  $GT=0.07$ . The sensor input is  $\theta_a$  in degrees and the output is in volts.

#### V. PLANT

The control system of the robot arm has been shown in Fig. 1. This system shown the sampling time,  $T = 0.1s$ , the power amplifier increase,  $K = 2.4$  and the sensor feedback gain,  $H_k = 0.07$ . the system phase margin with,  $D(z) = 1$ . the ZOH-TF can be define as

$$G_{HO}(s) = \frac{1-e^{-sT}}{s} \quad (7)$$

The plant TF in continuous time

$$G_p(s) = \frac{9.6}{s^2+2s} \quad (8)$$

The sensor gain feedback TF is used in a continuous-time plant

$$G_c(s) = G_p(s) \times H_k = \frac{0.672}{s^2+2s} \quad (9)$$

Where the TF is transfer function.

The sensor gain feedback TF that operates in discrete time is known as a discrete-time plant

$$G_d = \frac{0.00028289 (s+3.39e04)}{(s+1.524) (s+0.4406)} \quad (10)$$

Fig. 2. Introduces  $D(z) = 1$  system bode diagram,  $Pm$ , for the uncompensated system is  $79.6 \text{ deg}$ . With a gain margin  $Gm = 35.8 \text{ dB}$ .

#### A. Design of Phase Lag Controller

design a phase-lag controller with a dc gain of 10 that yields a system phase margin

$$G_{hf}(dB) = 20 \log \frac{a_0 \omega_{wp}}{\omega_{w0}} \quad (11)$$

The controller in this paper is built for  $39.7842$  degrees. For this design, the  $Pm$  and the cross-over or  $Pm$  frequency have been chosen as  $\omega_{wc} = 1.1291 \text{ rad}^{-1}$ .

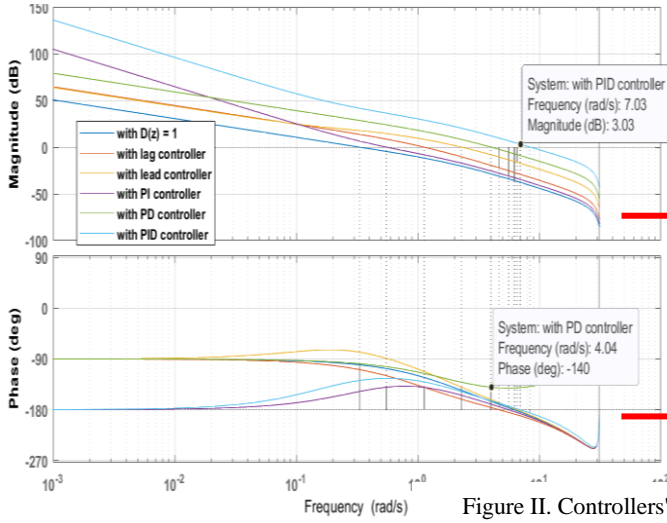
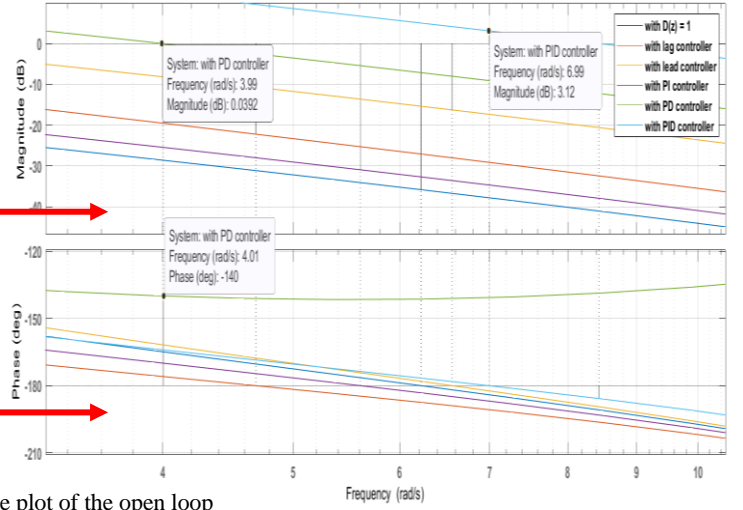


Figure II. Controllers' Bode plot of the open loop



$$\omega_{w0} = 0.1\omega_{wc} \quad (12)$$

and

$$\omega_{wp} = \frac{\omega_{w0}}{a_0|G_d(j\omega_{wc})|} \quad (13)$$

The TF of controller is

$$D_{lag}(z) = \frac{0.01167(z-0.7105)}{(z-0.9919)} \quad (14)$$

Figure II show the phase-lag controller. The compensated plant  $P_m$ ,  $P_m = 39.7842 \text{ deg}$  at  $0.0925 \text{ rad/s}$ , can be seen on the bode plot.  $G_m = 29.6 \text{ dB}$  at  $0.575 \text{ rad/s}$ . The gain and phase margin values are unknown in the marginalized bode plot of the controller, and hence these are determined to be infinite. The gain and phase margin values are unknown in the Bode plot of the controller, and hence these are found to be infinite. Where the  $P_m$  is the phase margin.

### B. Design of Phase Lead Controller

the phase-lead controller,  $a_0 = 10$  and maximum phase shift,  $\theta_m$  occurs at a frequency  $\omega_{wm} = \sqrt{\omega_{w0}\omega_{wp}}$ . The controller in this paper is equipped for 39.8827 degrees for this design, the  $P_m$  and cross over or  $P_m$  frequency have been chosen as  $2.2950 \text{ rad/s}$ . a phase-lead design controller with a dc gain of 10 that yields a system phase margin 40 deg.

$$D(j\omega_{wc})G_d(j\omega_{wc}) = 1\angle(180 + \phi_{pm}) \quad (15)$$

here  $\phi_{pm}$  pm is the desired  $P_m$  and

$$D(w) = a_0 \frac{1+w/(a_0/a_1)}{1+w/(b_1)^{-1}} \quad (16)$$

Where  $\omega_{w0} = a_0/a_1$  and  $\omega_{wp} = 1/b_1$ . The angle can be described as.

$$\theta_r = \angle D(j\omega_{wc}) = 180 + \phi_{pm} - \angle G_d(j\omega_{wc}) \quad (17)$$

The controller design requires

$$|D(j\omega_{wc})| = \frac{1}{|G_d(j\omega_{wc})|} \quad (18)$$

from the equation (16)-(18) it can be evaluated that

$$a_1 = \frac{1-a_0|G_d(j\omega_{wc})|\cos \theta_r}{\omega_{wc}|G_d(j\omega_{wc})\sin \theta_r} \quad (19)$$

and

$$b_1 = \frac{\cos \theta_r - a_0|G_d(j\omega_{wc})|}{\omega_{wc}\sin \theta_r} \quad (20)$$

Because of the phase lead characteristic,  $\theta_r > 0$  and in the design procedure  $\omega_{wc}$  has been constrained by the following requirements

$$\begin{aligned} \angle G_d(j\omega_{wc}) &< 180 + \phi_{pm}; |D(j\omega_{wc})| > a_0 \\ |G_d(j\omega_{wc})| &< \frac{1}{a_0}; b_1 > 0 \\ \cos \theta_c &> a_0|G_d(j\omega_{wc})| \end{aligned} \quad (21)$$

The transfer function of controller is

$$D(z) = \frac{10.424*(z-0.9832)}{(z-0.9618)} \quad (22)$$

From the bode plot, it can be observed that the compensated plant  $P_m$ ,  $P_m = 39.8827 \text{ deg}$ . At  $2.29 \text{ rad/s}$  and the gain margin,  $G_m = 16.2 \text{ dB}$  at  $6.57 \text{ rad/s}$ . From the bode plot of controller figure, in fact, it is obvious that the gain and  $P_m$  values are undefined and thereby these are found to be infinite.

### C. Design of PI Controller

PI controller means Proportional Integral controller it is composite of proportional and integral controller. They are in cascade with each other, as we see in fig.

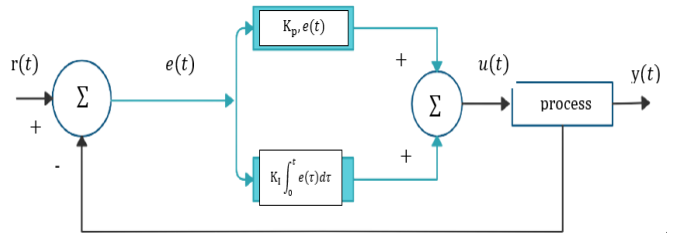


Figure III. THE DESIGN OF PI CONTROLLER

The TF of controller can be expressed as

$$D(w) = K_p + \frac{K_I}{w} = K_I \frac{1+w/\omega_{w0}}{w} \quad (23)$$

Table I. Controllers' Bode plot characteristics

Characteristics	Pm with D(z) = 1	Lag	Lead	PI	PD	PID
Gain Margin	[61.5755 1.7929e+04]	[12.8749 6.7538e+03]	[6.4935 1.7014e+03]	[0 9.9702 1.8455e+03]	Inf	[0 9.9702 1.8455e+03]
GM Frequency	[6.2240 31.4159]	[4.6917 31.4159]	[6.5659 31.4159]	[0 6.9890 31.4159]	Inf	[0 6.9890 31.4159]
Phase Margin	79.6399	39.7842	39.8827	38.1362	40.0	38.1362
PM Frequency	0.3315	1.1291	2.2950	1.8486	4.0035	1.8486
Delay Margin	41.9333	6.1499	3.0331	3.6006	1.7438	3.6006
DM Frequency	0.3315	1.1291	2.2950	1.8486	4.0035	1.8486
Stable	1	1	1	1	1	1

Where  $\omega_{w0} = K_I/K_P$ . However,

the discrete TF of a PI controller can be expressed as

$$D(z) = K_P + K_I \frac{Tz+1}{2z-1} \quad (24)$$

PI controller design that yields a system phase margin with 40 deg

$$D(j\omega_{wc})G_d(j\omega_{wc}) = 1\angle(-180 + \phi_{pm}) \quad (25)$$

Let  $A = |G_d(j\omega_{wc})|$ , The  $K_P$  proportional gain and  $K_I$  integral gain can be expressed as

$$K_P = \frac{\cos \theta_r}{|G_d(j\omega_{wc})|}$$

$$K_I = \frac{\omega_{wc} \sin \theta_r}{|G_d(j\omega_{wc})|}$$

The TF of controller is

$$D(z) = \frac{1.4575(z-0.9648)}{(z-1)} \quad (26)$$

From the bode plot, this is worth noting as that the compensated plant Pm,  $P_m = 40.0182$  deg. At  $0.553$  rad/s and the gain margin,  $G_m = 30.9$  dB at  $5.61$  rad/s. From the bode plot of controller figure, it can be observed that the gain and Pm values are undefined and thereby these are found to be infinite.

#### D. Design of PD Controller

PD controller means proportional derivative Controller so it has both, the proportional controller and derivative controller in cascade, so we have to add both, as we see in fig.

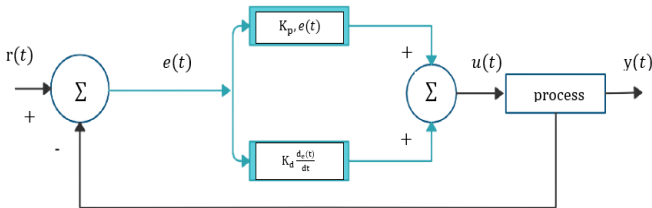


Figure IV. THE DESIGN OF PD CONTROLLER

Let  $w_1$  the Gain crssover frequency of the system with cascade PD controller,  $A_1$  the  $|G_1(j\omega)|$  at  $\omega_1$ . The PI controller transfer function is

$$D(\omega) = k_p + k_D \omega \quad (27)$$

and we can change the transfer function from w-plane to z - transform by using bilinear transformation. In Bilinear transformation  $\omega$  is replaced by  $\frac{2}{T} \left( \frac{z-1}{z+1} \right)$ , where  $T$  = sampling time.

the discrete TF of a PI controller is

$$D(z) = kp + k_D T^2 \left( \frac{z-1}{z+1} \right) \quad (28)$$

controller design that yields a system phase margin with 40 deg

$$-k_d \omega_1^2 + k_i + jk_p \omega_1 = \frac{\omega_1}{A_1} \sin \theta + j \frac{\omega_1}{A_1} \cos \theta \quad (29)$$

The derivative gain and proportional gain can be expressed as

$$k_d = \frac{\sin \theta}{\omega, A_1}$$

$$\therefore k_d = \frac{\sin \theta}{\omega_1 |G(j\omega_1)|} \quad (30)$$

$$k_p = \frac{\cos \theta}{A_1}$$

$$\therefore k_p = \frac{\cos \theta}{|G(j\omega_1)|} \quad (31)$$

#### E. Design of PID Controller

PID controller proportional Integral Derivative controller It consist of proportional, integral and Derivative controller all connected in the form of cascade, as we see in fig.

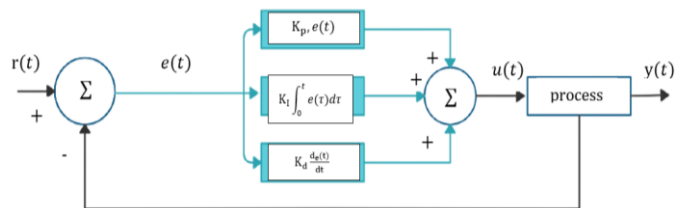


Figure V. THE DESIGN OF PID CONTROLLER

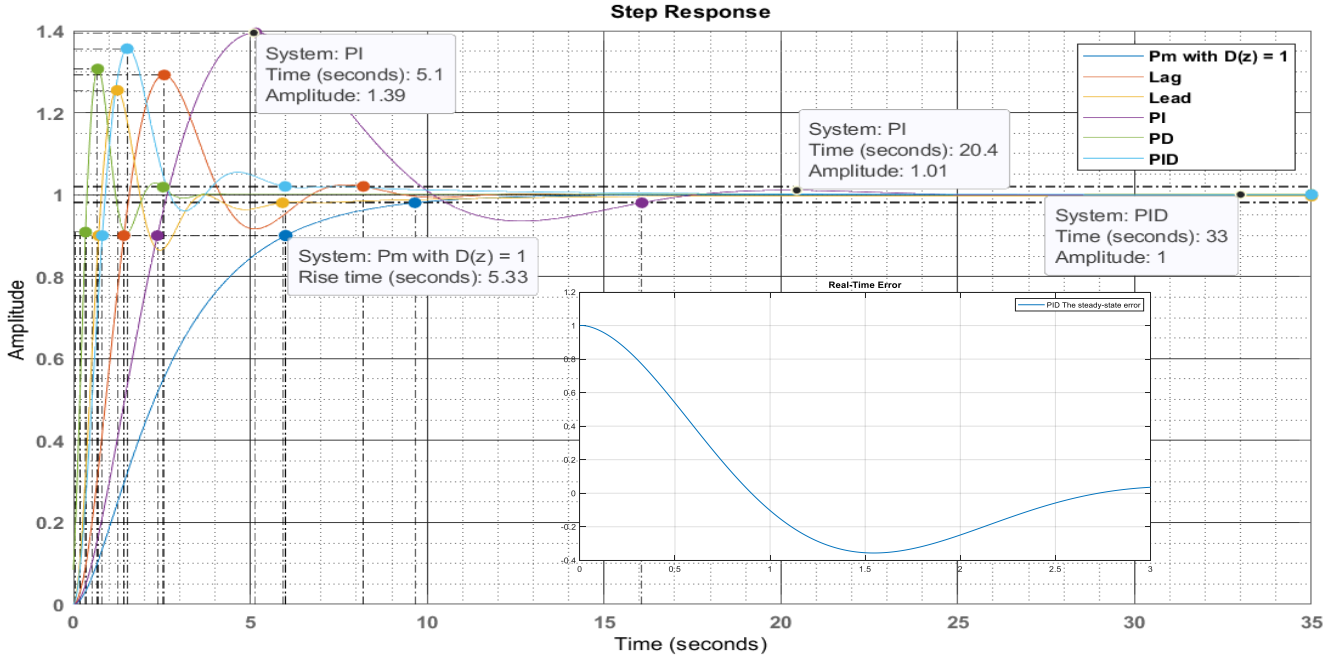


Figure VI. Controllers' step response of the closed loop

Table II. Controllers' step response characteristics

Characteristics	Pm with D(z) = 1	Lag	Lead	PI	PD	PID
Rise Time:	5.3278	1.0700	0.5359	1.8526	0.2940	0.6182
Settling Time	9.6546	8.1969	5.9127	16.0705	2.5379	5.9852
Settling Min	0.9020	0.9004	0.8641	0.9053	0.9068	0.9321
Settling Max	0.9986	1.2919	1.2543	1.3949	1.3064	1.3560
Overshoot	0	29.1881	25.4273	39.4865	30.6421	35.6044
Undershoot	0	0	0	0	0	0
Peak	0.9986	1.2919	1.2543	1.3949	1.3064	1.3560
Peak Time	15.6573	2.5692	1.2423	5.1509	0.6846	1.5198

$$D(w) = K_p + \frac{K_I}{w} + K_D w \quad (32)$$

and we can change the transfer function from  $w$ -plane to  $z$ -transform by using bilinear transformation. In Bilinear transformation  $\omega$  is replaced by  $\frac{2}{T} \left( \frac{z-1}{z+1} \right)$ , where  $T$  = sampling time.

The PID controller's discrete TF can be expressed as

$$D(z) = K_p + K_I \frac{T}{2} \frac{z+1}{z-1} + K_D \frac{z-1}{Tz} \quad (33)$$

controller design that yields a system phase margin with 40 deg

$$\left[ K_p + \frac{K_D \omega_{wc}^2 (2/T)}{(2/T)^2 + \omega_{wc}^2} \right] + j \left[ \frac{K_D \omega_{wc} (2/T)^2}{(2/T)^2 + \omega_{wc}^2} - \frac{K_I}{\omega_{wc}} \right] = K_R + jK_C$$

The  $K_p$  proportional gain,  $K_D$  derivative gain and  $K_I$  integral

gain can be expressed as the controller transfer function, which can be expressed as

$$\therefore K_p + \frac{K_D \omega_{wc}^2 (2/T)}{(2/T)^2 + \omega_{wc}^2} = \frac{\cos \theta_r}{|G_d(j\omega_{wc})|} \quad (34)$$

$$\therefore \frac{K_D \omega_{wc} (2/T)^2}{(2/T)^2 + \omega_{wc}^2} - \frac{K_I}{\omega_{wc}} = \frac{\sin \theta_r}{|G_d(j\omega_{wc})|} \quad (35)$$

The gain margin is = 20 dB at 6.99 rad/s, and the Pm is = 38.1 degrees by the PID controller at 1.85 rad/s. The gain and phase margin values are unknown in the marginalized bode plot of the controller, and hence these are determined to be infinite.

## VI. Step Response Characteristics

design problem explained in this paper has assumed an input of  $\theta_c = 0.07u(t)$ . The controllers scaled step response of the closed loop system for the designed is presented Figure VI. For the step response overshoot,  $\xi \downarrow \Rightarrow M_p \% (\%OS) \uparrow$ . From

the figure VI. And table II the best one is PID controller because it very less steady state error compared to other controllers. For PID  $\rightarrow$  steady state error  $\propto \frac{1}{k_p}$ , OS  $\propto \frac{1}{k_d}$ , Rise Time  $\propto \frac{1}{k_p}$ ,  $\propto \frac{1}{k_i}$ , Settling Time  $\propto \frac{1}{k_d}$ .  $K_p, K_d$ , and  $K_i$  can be described as the proportional, derivative, and integral parameters. The closed loop controls system is affected by all three of these parameters. In addition to those factors, the slow rising, slow settling, and long overshoot as well as the steady state error are also affected. A lag compensator shifts the Bode magnitude plot down at mid and high frequencies with its attenuation property. highlights for specification on steady-state error, the low frequency gain is changed. The proportional integral controller is equivalent to a control system that produces an output, this calls attention to which is the result of adding outputs from the proportional and integral controllers. PID is used in systems where proportional, integral, and derivative controllers are in use to compute an output. implies It's also there to reduce steady state error and improve stability. implies It's also there to reduce steady state error and improve stability. reveals that when used in conjunction with a proportional and a derivative controller, the proportional derivative controller generates an output, which is the product of the proportional and derivative controllers. If PD is being used, noise may be suppressed in the higher frequencies.

## VII. conclusion

This paper examines the performance and design assessment of five simple digital controllers, including lag, lead, PI, PD, and PID controllers, which are used for a physical robot arm joint plant. implies design a system controller with a dc gain of 10 that yields a system phase The design methodologies have been investigated in both discrete z-domain time approaches and warped s-domain or w-plane time frames. The controllers have been simulated on

MATLAB and bode plots with open loop and closed loop step response curves have been analyzed for comparative margin Pm 40 deg. C-O frequency is a crucial design specification to compensate the plant. premises. the suggests tells us Such design crucial as its specifications are applicable in different practical control systems.

## References

- [1] Chowdhury, Dhiman. "Design and Performance Analysis of Digital Controllers in Discrete and Continuous Time Domains for a Robot Control System." *Global Journal of Research In Engineering* (2018).
- [2] Unglaub, Ricardo AG, and D. Chit-Sang Tsang. "Phase tracking by fuzzy control loop." 1999 IEEE Aerospace Conference. Proceedings (Cat. No. 99TH8403). Vol. 5. IEEE, 1999.
- [3] Mastinu, Gianpiero, and Manfred Plöchl, eds. *Road and off-road vehicle system dynamics handbook*. CRC press, 2014.
- [4] Chowdhury, Dhiman, and Mrinmoy Sarkar. "Digital Controllers in Discrete and Continuous Time Domains for a Robot Arm Manipulator." *arXiv preprint arXiv:1912.09020* (2019).
- [5] Alassar, Ahmed Z., Iyad M. Abuhadrous, and Hatem A. Elaydi. "Comparison between FLC and PID Controller for 5DOF robot arm." 2010 2nd International Conference on Advanced Computer Control. Vol. 5. IEEE, 2010.
- [6] Phillips, Charles L., and H. Troy Nagel. *Digital control system analysis and design*. Prentice-Hall, Inc., 1989.
- [7] Misir, Dave, Heidar A. Malki, and Guanrong Chen. "Design and analysis of a fuzzy proportional-integral-derivative controller." *Fuzzy sets and systems* 79.3 (1996): 297-314.
- [8] Klee, Harold, and Joe Dumas. "Theory, simulation, experimentation: an integrated approach to teaching digital control systems." *IEEE transactions on education* 37.1 (1994): 57-62.
- [9] Liu, Hui. *Robot Systems for Rail Transit Applications*. Elsevier, 2020.
- [10] Boukas EK., AL-Sunni F.M. (2011) Design Based on Transfer Function. In: *Mechatronic Systems*. Springer, Berlin, Heidelberg. [https://doi.org/10.1007/978-3-642-22324-2\\_5](https://doi.org/10.1007/978-3-642-22324-2_5).



# Machine learning model optimization with hyper-parameter tuning approach

Md Riyad Hossain  
Department of Manufacturing  
Engineering  
University of Texas Rio Grande Valley,  
USA  
md.hossain01@utrgv.edu

Douglas Timmer  
Department of Manufacturing  
Engineering  
University of Texas Rio Grande Valley,  
USA  
douglas.timmer@utrgv.edu

Hiram Moya  
Department of Manufacturing Engineering  
University of Texas Rio Grande Valley,  
USA  
hiram.moya@utrgv.edu

**Abstract**— Hyper-parameters tuning is a key step to find the optimal machine learning parameters. Determining the best hyper-parameters takes a good deal of time, especially when the objective functions are costly to determine, or myriad parameters are required to be tuned. In contrast to the conventional machine learning algorithms, Neural Network requires tuning hyper-parameters more because it must process a lot of parameters together, and depending on the fine tuning, the accuracy of the model can be varied in between 25%-90%. A few of the most effective techniques for tuning hyper-parameters in the Deep learning methods are: Grid search, Random forest, Bayesian optimization, etc. Every method has some advantages and disadvantages over others. For example: Grid search has proven to be an effective technique to tune hyper-parameters despite some drawbacks like trying too many combinations and performing poorly in case of tuning many parameters simultaneously. In our work, we will determine, show and analyze the efficiencies of a real-world synthetic polymer dataset for different parameters and tuning methods.

**Keywords**— Machine learning, Hyperparameter optimization, Grid Search technique, Random Search, BO-GP

## I. Introduction

In the era of Machine learning, performance (based on accuracy and computing time) is very important. The growing number of tuning parameters associated with the Machine learning models is tedious and time-consuming to set by standard optimization techniques. Researchers working with ML models often spend long hours to determine the perfect combination of hyper-parameters [1]. If we think  $w$ ,  $x$ ,  $y$ ,  $z$  as the parameters of the model, and if all of these parameters are integers ranging from 0.0001 to say 5.00, then hyperparameter tuning is the finding the best combinations to make the objective function optimal.

One of the major difficulties in working with the Machine learning problem is tuning hyperparameters. These are the design parameters that could directly affect the training outcome. The conversion from a non-tuned Machine learning model to a tuned ML model is like learning to predict everything accurately from predicting nothing correctly [2]. There are two types of parameters in ML models: Hyperparameters, and Model parameters. Hyperparameters are arbitrarily set by the user even before starting to train the model, whereas, the model parameters are learned during the training.

The quality of a predictive model mostly depends on the configuration of its hyperparameters, but it is often difficult to know how these hyperparameters interact with each other to affect the final results of the model [14]. To determine accuracy and make a comparison between two models it is always better to make comparisons between two models with

both of the models' parameters tuned. It is not expected to compare a Decision Tree model with the already tuned parameter versus an ANN model whose hyperparameters haven't been optimized yet.

## II. Literature Review

The hyperparameter tuning, due to its importance, has changed to a new interesting topic in the ML community. The hyperparameter tuning algorithms are either model-free or model-based. Model-free algorithms are free of using knowledge about the solution space extracted during the optimization; a few of this category includes manual search [4], random search [2, 6-7], and grid search [5]. In the Manual search categories, we assume the values of the parameters from our previous experience. In this technique, the user allows to set hyperparameters values based on judgments or previous experience, trains the algorithm by them, observes the performance, keeps doing it to train the model until achieving a standard accuracy and then selects the best set of hyperparameters that gives the maximum accuracy. However, this technique is heavily dependent on the judgment and previous expertise and its reliability is dependent on the correctness of the previous knowledge [3]. Some of the few of the main parameters used by Random forest classifiers are criterion, max\_depth, n\_estimators, min\_samples\_split etc.

In the Random search, we train and test our model based on some random combinations of the hyperparameters. This method is better used to identify new combinations of the parameters or to discover new hyperparameters. Although it may take more time to process, it often leads to better performance. Bergstra et al. (2012) in their work mentioned that, over the same domain, random search can find models that are as good as or even better in a reduced computation time. After granting the same budget in terms of computational constraints for the random search, it was evident that random search can deliver best models within a larger and reduced promising configuration spaces [16]. Random Search, which is developed based on grid research, examine a set of random combinations to develop and train the algorithm; Bergstra et al. (2011) [2].

In the grid search, the user sets a matrix of hyperparameters and trains the model based on each possible combination. Amirabadi et al. (2020) proposes two novel suboptimal grid search techniques on the four separate datasets to show the

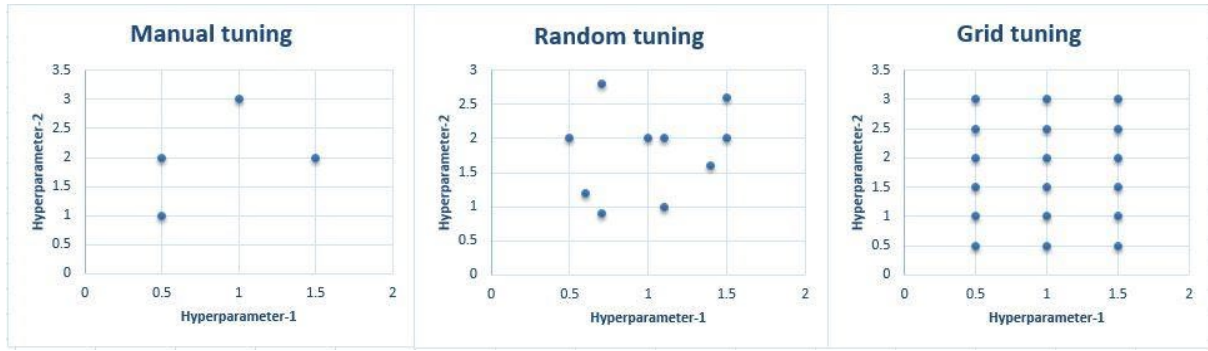


Figure I. (a) Manual tuning (b) Random tuning (c) Grid tuning approach [From left to Right]

efficiency of their hyperparameter tuning model and later compare it with some of the other recently published work. The main drawback of the grid search method is its high complexity. It is commonly used when there are a few numbers of hyperparameters to be tuned. In other words, grid search works well when the best combinations are already determined. Some of the similar works of grid search applications have been reported by Zhang et al. (2014) [17], Ghawi et al. (2019) [18], and Beyramysoltan et al. (2013) [19].

Zhang et al. (2019) [20] in their work reported a few of the drawbacks of the existing hyperparameter tuning methods. In their work, they mentioned grid search as an ad-hoc process, as it traverses all the possible combinations, and the entire procedure requires a lot of time. Andradóttir (2014) [13] shows that Random Search (RS) eradicates some of the limitations of the grid search technique to an extent. RS can reduce the overall time consumption, but the main disadvantage is that it cannot converge to the global optimal value.

The combination of randomly selected hyper-parameters can never guarantee a steady and widely acceptable result. That's why, apart from the manually tuning methods, automated tuning methods are becoming more and more popular in recent times; snoek et al. (2015) [10]. Bayesian Optimization is one of the most widely used automated hyperparameter tuning methods to find the global optimum in fewer steps. However, Bayesian optimization's results are sensitive to the parameters of the surrogate model and the accuracy is greatly depending on the quality of the learning model; Amirabadi et al. (2020) [3].

To minimize the error function of hyperparameter values, Bayesian optimization adopts probabilistic surrogate models like Gaussian processes. Through precise exploration and development, an alternative model of hyperparameter space is established; Eggenberger et al. (2013) [8]. However, probabilistic surrogates need accurate estimations of sufficient statistics of error function distribution. So, a sizable number of hyperparameters is required to evaluate the estimations and this method doesn't work well when there is to process myriad hyperparameters altogether.

### III. Methodology

The purpose of hyperparameter optimization is to find the global optimal value  $x^*$  of the objective function  $f(\mathbf{x})$  can be

evaluated for any arbitrary  $x \in X$ ,  $x^* = \arg \min_{x \in X} f(x)$ , and  $X$  is a hyperparameter space that can contain categorical, discrete, and continuous variables [27]. In order to construct the design of different machine learning models, the application of effective hyperparameter optimization techniques can simplify the process of identifying the best hyperparameters for the models. HPO contains four major components: First, an estimator that could be a regressor or any classifier with one or more objective functions, second: a search space, Third: an optimization method to find the best combinations, and Fourth: a function to make a comparison between the effectiveness of various hyperparameter configurations [28].

#### A. Grid Search

Grid search is a process that exhaustively searches a manually specified subset of the hyperparameter space of the target algorithm [30]. A traditional approach to finding the optimum is to do a grid search, for example, to run experiments or processes on a number of conditions, for example, if there are three factors, a  $15 \times 15 \times 15$  would mean performing 3375 experiments under different conditions. [32]. Grid search is more practical when [31]: (1) the total number of parameters in the model is small, say  $M < 10$ . The grid is  $M$ -dimensional, so the number of test solutions is proportional to  $L^M$ , where  $L$  is the number of test solutions along each dimension of the grid. (2) The solution is known to be within a specific range of values, which can be used to define the limits of the grid. (3) The direct problem  $d = g(m)$  can be computed quickly enough that the time required to compute  $L^M$  from them is not prohibitive. (4) The error function  $E(m)$  is uniform on the scale of the grid spacing,  $\Delta m$ , so that the minimum is not lost because the grid spacing is too coarse.

There are many problems with the grid search method. The first is that the number of experiments can be prohibitive if there are several factors. The second is that there can be significant experimental error, which means that if the experiments are repeated under identical conditions, different responses can be obtained; therefore, choosing the best point on the grid can be misleading, especially if the optimum is fairly flat. The third is that the initial grid may be too small for the number of experiments to be feasible, and it could lose characteristics close to the optimum or find a false (local) optimum [32].

## B. Random Search

Random search [33] is a basic improvement on grid search. It started with a randomized search over hyperparameters from certain distributions over approximate parameter values. This searching process runs as long as the predetermined budget is exhausted, or at least until achieving a desired set of accuracy. These methods are the simplest stochastic optimization and are very useful for certain problems, such as small search space and fast-running simulation. RS finds a value for each hyperparameter, prior to the probability distribution function. Both the GS and RS estimate the cost measure based on the produced hyperparameter sets. Although RS is simple, it has proven to be more effective than Grid search in many of the cases [33]. Random search has been shown to provide better results due to several benefits: first, the budget can be set independently based on the distribution of the search space, therefore, random search technique can sometime work better especially if the multiple hyper-parameters are not uniformly distributed [34]. Second: Because each evaluation is independent, it is easy to parallelize and allocate resources. Unlike GS, RS samples a few parameter combinations from a defined distribution, which maximizes system efficiency by reducing the likelihood of wasting a lot of time in a small, underperforming area. In addition, this method can detect global optimum values or close to global if given a sufficient budget. Third, although getting optimal outputs applying random search is not promising, lengthy processing time may lead to a greater likelihood of getting the best hyperparameter set, whereas extra search times cannot always guarantee improved results in Grid searches. The use of random search is recommended in the initial stages of HPO to narrow the search space quickly, before using guided algorithms to get better results. The main drawback [28] of RS and GS is that each evaluation in its iteration does not depend on previous evaluations; thus, they waste time evaluating underperforming areas of the search space.

## C. Bayesian Optimization

Bayesian optimization (BO) is a commonly used reprocessing algorithm for HPO problems. Unlike GS and RS, BO determines future assessment levels based on the previous results. To determine the following parameters of the hyperparameter, BO uses two key factors: a surrogate model and an acquisition function. The division model aims to match all the points that are now seen in the objective function. The acquisition function determines the use of different points, balancing exploration and exploitation. The BO model balances the search and use process to identify the best possible area and avoid losing the best configuration in undeveloped areas [35].

The basic BO method works as follows: (i) Building a reduced-order probabilistic model (ROM) of the objective function. (ii) Finding the best hyperparameter values in the ROM model. (iii) Applying those optimal values to the objective function. (iv) Updating the ROM model with the new set of results. (v) Repeating above steps until achieving maximum number of iterations.

BO is more efficient than GS and RS because it can detect optimal combinations of hyperparameters by analyzing previously tested values, and running the surrogate model is

usually much cheaper than running the objective function. However, because Bayesian optimization models are run based on previously tested values, it is difficult to belong to them with parallel sequential methods; but they are generally able to detect optimal close hyperparameter combinations in a few iterations [36]. Common substitution models for BO include the Gaussian process (GP) [37], random forest (RF) [38]. Therefore, there are three main BO algorithms based on their substitution models: BO-GP, BO-RF, BO-TPE. GP is an attractive reduced order model of BO that can be used to quantify forecast uncertainty. This is not a parametric model and the number of its parameters depends only on the input points. With the right kernel function, your GP can take advantage of the data structure. However, the GP also has disadvantages. For example, it is conceptually difficult to understand with BO theory. In addition, its low scalability with large dimensions or many data points is another important issue [36].

## IV. Dataset description & Basics of polymer extrusion

### A. Denier

Denier is a weight measurement usually refers to the thickness of the threads. It is the weight (grams) of a single optical fiber for 9 kilometers. If we have a 9 km fiber weighs 1 gram, this fiber has a denier of 1, or 1D. A fiber with less than 1-gram weight calls Microfibers [22]. Microfibers become a new development trend in the synthetic polymer industry. The higher the denier is, the thicker and stronger the fiber is. Conversely, less denier means that the fiber/fabric will be softer and more transparent. Fine denier fibers are becoming a new standard and are very useful for the development of new textiles with excellent performance [21].

### B. Breaking Elongation (%)

Elongation at break is one of the few main quality parameters of any synthetic fiber [24]. It is the percentage of elongation at break. Fiber elongation partly reflects the extent of stretching a filament under a certain loading condition. Fibers with high elongation at break are determined to be easily stretched under a predetermined load. Fibers showing these characteristics are known to be flexible. The elongation behavior of any single fiber can be complex because of its multiplicity of structural factors affecting it. Moreover, a cotton fiber comes up with a natural crimp, which is important for fibers to stick together while undergoing other production processes [23]. If  $L$  is the length of the fiber, then the equation for the percentages of the breaking elongation would be:

$$E_{\text{Breaking elongation}} = \frac{\Delta L_{\text{Break}}}{L_0} * 100\%$$

Breaking elongation for the cotton fiber might be varied from 5% to 10%, which is significantly lower than that of wool fibers (25%-45%), and much lower than polyester fibers (typically over 50%).

### C. Breaking force (cN) and Tenacity (cN/tex)

Breaking tenacity is the maximum load that a single fiber can withstand before breaking. For the Polypropylene and

PET staple fibers, 10 mm lengths sample filaments are drawn until failure. Breaking tenacity is measured in grams/denier. Very small forces are encountered when evaluating fiber properties, so an instrument with gram-level accuracy is required [25]. The tenacity of virgin PP fibers is about 5–8 g/den, and the elongation at break is about 100%. At the same time, the tenacity of recycled PET is about 3.5-5.7 g/den; the elongation at break usually exceeds 100%.

#### D. Draw Ratio

The drawing ration is the ratio of the diameter of the initial blank form to the diameter of the drawn part. The limiting drawing ratio (Capstan speed/Nip reel speed) for the extruder section is between 1.6 and 2.2 [26], whereas, for the stretching section it is in between 3 and 4.

### V. Results

#### A. MSE for default hyperparameters

We ran the code in Google Colaboratory and got an MSE value of 44.8%, 3653.6%, 3100.7%, and 713.7% accordingly for Random Forest, Support Vector Machine, K-Nearest Neighbors, and Artificial Neural Network.

Table I. MSE & Cycle time for Default hyperparameters

Name	MSE	Cycle time (s)
RF	0.448	0.5
SVM	36.536	1
KNN	31.007	0.8
ANN	7.137	5

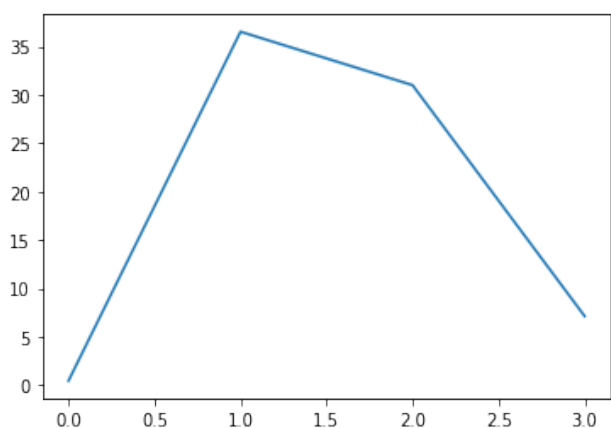


Figure II. Graphical representations of MSE values for default hyperparameters

#### B. MSE for Grid Search

The MSE value of 105.3%, 92.7%, 2945.7%, and 47.5% accordingly for Random Forest, Support Vector Machine, K-Nearest Neighbors, and Artificial Neural Network. This result is much better and improved compared to default hyperparameter values.

Table II. MSE & Cycle time for Grid Search technique

Name	MSE	Cycle time (s)
RF	1.053	5
SVM	0.927	2
KNN	29.45	1
ANN	0.475	5

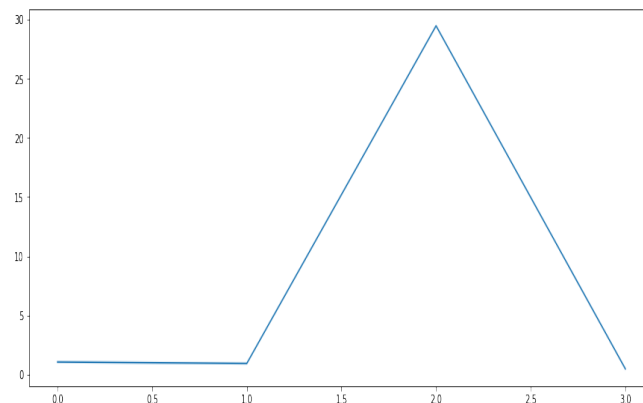


Figure III. Graphical representations of MSE values for Grid Search

#### C. MSE for Random Search

The MSE value of 324.3%, 275.1%, 2945.5%, and 30.4% accordingly for Random Forest, Support Vector Machine, K-Nearest Neighbors, and Artificial Neural Network.

Table III. MSE & Cycle time for Grid Search technique

Name	MSE	Cycle time (s)
RF	3.243	3
SVM	2.751	1
KNN	29.45	2
ANN	0.304	5

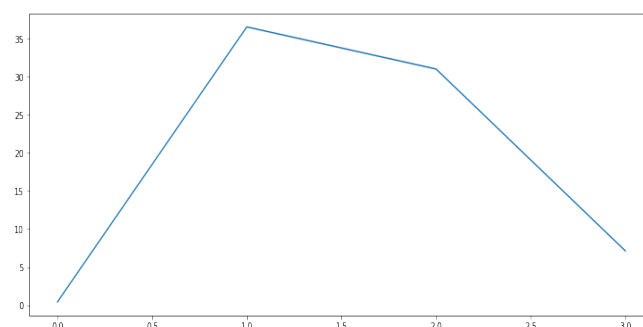


Figure IV. Graphical representations of MSE values for Random Search

### C. MSE for Bayesian Optimization with Gaussian Process (BO-GP)

The MSE value for BO-GP are 723.5%, 648.5%, 2945%, and 30.8% accordingly for Random Forest, Support Vector Machine, K-Nearest Neighbors, and Artificial Neural Network.

Table IV. MSE & Cycle time for BO-GP Search technique

Name	MSE	Cycle time (s)
RF	7.235	2
SVM	6.485	3
KNN	29.45	2
ANN	0.308	5

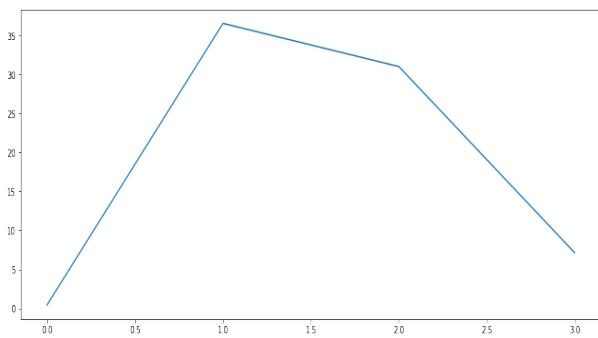


Figure V. Graphical representations of MSE values for BO-GP

### VI. Discussion

In order to put the theory into practice, several experiments have been performed on an industrial-based synthetic polymer model. This section describes experiments with four different HPO techniques on three general and representative ML algorithms. In the first part of the section, we discussed the experimental setup and the main HPO process. In the second part, we compare and analyze the results of the application of different HPO methods.

Table V. Performance evaluation of applying HPO methods to the regressor on the synthetic polymer dataset

Model	Optimization Algorithm	MSE	Cycle time (s)
RF	Default HP's	0.448	0.5
	GS	1.053	5
	RS	3.243	3
	BO-GP	7.235	2
SVM	Default HP's	36.536	1
	GS	0.927	2
	RS	2.751	1
	BO-GP	6.485	3
KNN	Default HP's	31.007	0.8
	GS	29.45	1

	RS	29.45	2
	BO-GP	29.45	2
ANN	Default HP's	7.137	5
	GS	0.475	5
	RS	0.304	5
	BO-GP	0.308	5

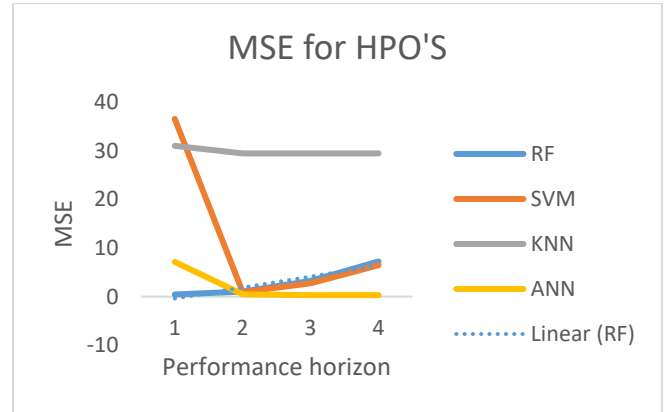


Figure VI. Graphical representations of MSE for different HPO's

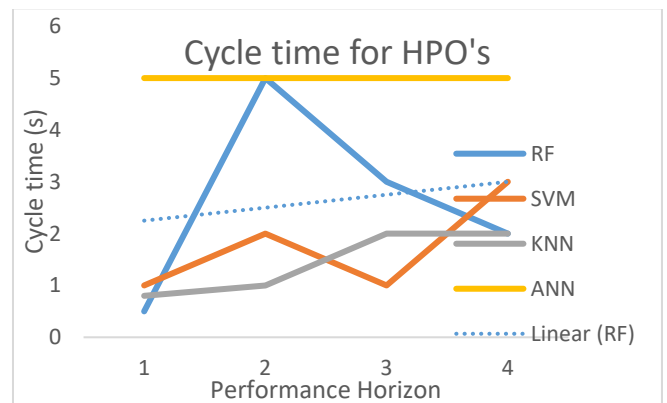


Figure VII. Graphical representations of Cycle time for different HPO's

### VI. Conclusion

Machine learning has become the primary strategy for dealing with data problems and is widely used in various applications. To apply ML models to practical problems, hyperparameters must be tuned to handle specific datasets. However, as the size of the generated data increases greatly in real life, and manual tuning of hyperparameters is extremely computationally expensive, it has become essential to optimize the hyperparameters by an automatic process. In this work, we used hyperparameter techniques in the ML model to find the best set of hyperparameters. Our data set was small, and in this small dataset we can see that the randomly selected subsets are very representative for the given data set, as they can effectively optimize all types of hyperparameters. Our future work would be to test our dataset with an advanced Machine learning technique probably with deep learning models and see the feedback.

## References

- [1] Cho, H., Kim, Y., Lee, E., Choi, D., Lee, Y., & Rhee, W. (2020). Basic Enhancement Strategies When Using Bayesian Optimization for Hyperparameter Tuning of Deep Neural Networks. *IEEE Access*, 8, 52588–52608. doi:10.1109/access.2020.2981072
- [2] J.S. Bergstra, R. Bardenet, Y. Bengio, B. Kégl, Algorithms for hyperparameter optimization, in: *Advances in Neural Information Processing Systems*, 2011, pp. 2546–2554.
- [3] Amirabadi, M., Kahaei, M., & Nezamalhosseini, S. (2020). Novel suboptimal approaches for hyperparameter tuning of deep neural network [under the shelf of optical communication]. *Physical Communication*, 41, 101057. doi:10.1016/j.phycom.2020.101057
- [4] F. Hutter, J. Lücke, L. Schmidt-Thieme, Beyond manual tuning of hyperparameters, *DISKI 29 (4)* (2015) 329–337.
- [5] F. Friedrichs, C. Igel, Evolutionary tuning of multiple SVM parameters, *Neurocomputing* 64 (2005) 107–117.
- [6] R.G. Mantovani, A.L. Rossi, J. Vanschoren, B. Bischl, A.C. De Carvalho, Effectiveness of random search in SVM hyperparameter tuning, in: *2015 International Joint Conference on Neural Networks (IJCNN)*, 2015, pp. 1–8.
- [7] L. Li, A. Talwalkar, Random search and reproducibility for neural architecture search, 2019, arXiv preprint arXiv:1902.07638.
- [8] K. Eggenberger, M. Feurer, F. Hutter, J. Bergstra, J. Snoek, H. Hoos, K. Leyton-Brown, Towards an empirical foundation for assessing bayesian optimization of hyperparameters. In *NIPS workshop on Bayesian Optimization in Theory and Practice (Vol. 10, 3)*, 2013.
- [9] H. Larochelle, D. Erhan, A. Courville, J. Bergstra, Y. Bengio, An empirical evaluation of deep architectures on problems with many factors of variation, in: *Proceedings of the 24th International Conference on Machine Learning, ACM*, 2007, pp. 473–480.
- [10] J. Snoek, O. Rippel, K. Swersky, R. Kiros, N. Satish, N. Sundaram, et al., Scalable bayesian optimization using deep neural networks, in: *International conference on machine learning*, 2015, pp. 2171–2180.
- [11] Jones, D.R.; Schonlau, M.; Welch, W.J. Efficient global optimization of expensive black-box functions. *J. Glob. Optim.* 1998, 13, 455–492.
- [12] Calandra, R.; Peters, J.; Rasmussen, C.E.; Deisenroth, M.P. Manifold Gaussian processes for regression. In *Proceedings of the 2016 International Joint Conference on Neural Networks, Vancouver, BC, Canada, 24–29 July 2016*; pp. 3338–3345.
- [13] Andrad ottir, S.: A review of random search methods. In: *Handbook of Simulation Optimization*, pp. 277{292. Springer (2015)
- [14] Li, L.; Jamieson, K.; DeSalvo, G.; Rostamizadeh, A.; Talwalkar, A. Hyperband: A Novel Bandit-Based Approach to Hyperparameter Optimization. *Journal of Machine Learning Research* 18 (2018) 1-52
- [15] A. Klein, S. Falkner, J. T. Springenberg, and F. Hutter. Learning curve prediction with Bayesian neural networks. In *International Conference On Learning Representation (ICLR)*, 2017.
- [16] J.S. Bergstra, Y. Bengio. Random Search for Hyper-Parameter Optimization, in: *Journal of Machine Learning Research* 13 (2012) 281-305
- [17] Zhang, H., Chen, L., Qu, Y., Zhao, G., & Guo, Z. (2014). Support Vector Regression Based on Grid-Search Method for Short-Term Wind Power Forecasting. *Journal of Applied Mathematics*, 2014, 1-11. doi:10.1155/2014/835791
- [18] Ghawi, R., & Pfeffer, J. (2019). Efficient Hyperparameter Tuning with Grid Search for Text Categorization using kNN Approach with BM25 Similarity. *Open Computer Science*, 9(1), 160–180. doi:10.1515/comp-2019-0011
- [19] Beyramysoltan, S., Rajkó, R., & Abdollahi, H. (2013). Investigation of the equality constraint effect on the reduction of the rotational ambiguity in three-component system using a novel grid search method. *Analytica Chimica Acta*, 791, 25–35. Doi: 10.1016/j.aca.2013.06.043
- [20] Zhang, X., Chen, X., Yao, L., Ge, C., & Dong, M. (2019). Deep Neural Network Hyperparameter Optimization with Orthogonal Array Tuning. *Communications in Computer and Information Science Neural Information Processing*, 287-295. doi:10.1007/978-3-030-36808-1\_31
- [21] Zhang, C., Liu, Y., Liu, S. et al. Crystalline behaviors and phase transition during the manufacture of fine denier PA6 fibers. *Sci. China Ser. B-Chem.* 52, 1835 (2009). https://doi.org/10.1007/s11426-009-0242-5
- [22] Joe. (2020, May 5). What Is Denier Rating? Why Does It Matter To You? DigiTravelist. https://www.digitravelist.com/what-is-denier-rating/.
- [23] Elmogahzy, Yehia (2018). Handbook of Properties of Textile and Technical Fibres || Tensile properties of cotton fibers. , (), 223–273. doi:10.1016/B978-0-08-101272-7.00007-9
- [24] Tyagi, G.K. (2010). Advances in Yarn Spinning Technology || Yarn structure and properties from different spinning techniques. , (), 119–154. doi:10.1533/9780857090218.1.119
- [25] Blair, K. (2007). Materials and design for sports apparel. *Materials in Sports Equipment*, 60-86. doi:10.1533/9781845693664.1.60
- [26] Swift, K., & Booker, J. (2013). Forming Processes. *Manufacturing Process Selection Handbook*, 93-140. doi:10.1016/b978-0-08-099360-7.00004-5
- [27] Cho, H., Kim, Y., Lee, E., Choi, D., Lee, Y., & Rhee, W. (2020). Basic Enhancement Strategies When Using Bayesian Optimization for Hyperparameter Tuning of Deep Neural Networks. *IEEE Access*, 8, 52588–52608. doi:10.1109/access.2020.2981072
- [28] Yang, L., & Shami, A. (2020). On hyperparameter optimization of machine learning algorithms: Theory and practice. *Neurocomputing*, 415, 295-316. doi:10.1016/j.neucom.2020.07.061
- [29] Chan, S., & Treleven, P. (2015). Continuous Model Selection for Large-Scale Recommender Systems. *Handbook of Statistics Big Data Analytics*, 107-124. doi:10.1016/b978-0-444-63492-4.00005-8
- [30] Menke, W. (2012). Nonlinear Inverse Problems. *Geophysical Data Analysis: Discrete Inverse Theory*, 163-188. doi:10.1016/b978-0-12-397160-9.00009-6
- [31] Brereton, R. (2009). Steepest Ascent, Steepest Descent, and Gradient Methods. *Comprehensive Chemometrics*, 577-590. doi:10.1016/b978-044452701-1.00037-5
- [32] Bergstra, J., Bengio, Y., 2012. Random search for hyperparameter optimization. *J. Mach. Learn. Res.* 13, 281–305. ISSN 1532-4435. URL http://dl.acm.org/citation.cfm?id=2188385.2188395
- [33] Yu, T., & Zhu, H. (2020). Hyper-Parameter Optimization: A Review of Algorithms and Applications, https://arxiv.org/abs/2003.05689
- [34] E. Hazan, A. Klivans, and Y. Yuan, Hyperparameter optimization: a spectral approach, arXiv preprint arXiv:1706.00764, (2017). https://arxiv.org/abs/1706.00764.
- [35] Hutter, F., Kotthoff, L., & Vanschoren, J. (2019). *Automated Machine Learning Methods, Systems, Challenges*. Cham: Springer International Publishing.
- [36] Seeger, M. (2004). Gaussian Processes For Machine Learning. *International Journal of Neural Systems*, 14(02), 69-106. doi:10.1142/s0129065704001899
- [37] Hutter F., Hoos H.H., Leyton-Brown K. (2011) Sequential Model-Based Optimization for General Algorithm Configuration. In: Coello C.A.C. (eds) *Learning and Intelligent Optimization*. LION 2011. Lecture Notes in Computer Science, vol 6683. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-25566-3\_40
- [38] Bergstra J, Bardenet R, Bengio Y, Kégl B (2011) Algorithms for hyper-parameter optimization. *Adv Neural Inf Process Syst (NIPS)* 24:2546–2554

# Polyvinyl Alcohol/Cassava Starch/Nano-CaCO<sub>3</sub> based Nano-Biocomposite Films: Mechanical and Optical Properties

Eslem Kavas  
Department of Polymer Materials Engineering  
Bursa Technical University  
Bursa, Turkey  
eslemkavas123@gmail.com

Pinar Terzioglu\*  
Department of Polymer Materials Engineering  
Bursa Technical University  
Bursa, Turkey  
pinar.terzioglu@btu.edu.tr

**Abstract**— This paper reports the preparation and characterization of nano-CaCO<sub>3</sub> incorporated polyvinyl alcohol (PVA)/cassava starch biocomposite films. The films were prepared by solution casting method using a glass mold. The effect of nano-CaCO<sub>3</sub> content (1 and 2 %) on the structural, mechanical and optical properties of biocomposite films was investigated by Fourier transform infrared (FTIR) spectroscopy, universal testing machine and ultraviolet-visible-near infrared (UV-VIS-NIR) spectroscopy analyses, respectively. The mechanical tests revealed that the 1% nano-CaCO<sub>3</sub> made no changes on the tensile strength of the PVA/cassava starch films, while 2% nano-CaCO<sub>3</sub> addition decreased the tensile strength of the films. The transparency of films was slightly increased with nano-filler addition.

**Keywords**— *biocomposite, biobased polymers, calcium carbonate, nano filler*

## I. Introduction

The materials research in the packaging industry focuses on the use of biodegradable polymers to overcome the drawbacks of environmental problems caused by non-degradable plastics [1]. Among biodegradable polymers, PVA has been preferred in the commercial packaging industry own to its excellent film-forming capability, resistance to oil and mechanical properties [2],[3]. However, the blending of PVA with different polymers seems to be a good solution to its relatively high price [2]. Due to the wide availability in natural sources, biodegradability, low cost and non-toxicity, starch is usually used to blend with PVA [1],[2].

In the combination of PVA and starch that has excellent compatibility with each other, the final materials present physicochemical features at good levels as well as cost competitiveness [4],[5]. However, the investigations going on to obtain polymer materials with enhanced characteristics. The efficient method of widening functionality of PVA-starch materials is the incorporation of a small amount of a functional filler to the polymer matrix [6].

Nano-biocomposites are an important class of hybrid materials that can be obtained by incorporation of nano-sized filler (nanofiller) to a bio-based matrix [4]. The together use of eco-friendly polymers and nano-sized fillers, in order to obtain synergic effects, is considered as one of the most innovative way to improve the features of the polymer materials [4],[7]. Nanofillers can be classified into three different types as nano-fiber, nano-particle, nano-plate [8]. Among the nano-particles used for such applications, nano-calcium carbonate generally seems to be a good candidate which exhibited unique properties like being cost-effective, odorless, and high thermal stability [9],[10].

Previous studies that have used nano-CaCO<sub>3</sub> as a filler presented promising results to enhance the mechanical

[11],[12],[13],[14], oxygen barrier [12] and thermal properties [12] of the hybrid nano-biocomposite materials. Sudhir et al. prepared the rice starch/PVA/CaCO<sub>3</sub> composite films by solution casting method and investigated the effect of calcium carbonate amount on the fire retardant, tensile strength and thermal properties of nanobiocomposite film [12]. It was reported that the thermal and tensile properties of the composites increased with the increment of nano-CaCO<sub>3</sub> loading amount. The highest values were reached with 10 wt.% of CaCO<sub>3</sub> concentration. Furthermore, the oxygen permeation of composite films was reduced when the filler amount increased. Therefore, the developed material was suggested to be used in packaging applications. In another study, Fukuda et al. used calcium carbonate particles as inorganic fillers to enhance the properties of poly(l-lactide) (PLLA) composite films [13]. Higher results were obtained for Young's modulus of the 10 wt.% of the CaCO<sub>3</sub> particles incorporated films when compared with the pure PLLA film. Another work by Sun et al. [14] characterized the corn starch based films impregnated with nano-CaCO<sub>3</sub> particles. The results showed that the tensile strength of the films increased with the increased CaCO<sub>3</sub> content up to 0.06%. However, the 0.1% and 0.5% loading amount of CaCO<sub>3</sub> lead to decrease in the tensile strength of the films. The studies showed that the properties of the final composites can be varied mostly due to the polymer matrix composition, the filler-matrix interaction and the distribution of the filler.

Upon thorough researches, the addition of nano-CaCO<sub>3</sub> to PVA/cassava starch biocomposite film has not been reported to date. In the present work, it is proposed to evaluate the effect of nano-CaCO<sub>3</sub> content on the performance of PVA/cassava starch biocomposite films. The biocomposite films were compared for their structural, mechanical, and optical properties.

## II. Materials and Methods

### A. Materials

Polyvinyl alcohol that has 87.16% degree of hydrolysis and 95.4% purity was obtained from Zag Kimya, TURKEY. Cassava starch was purchased from Tito, TURKEY. Calcium carbonate (CaCO<sub>3</sub>) nanopowder was obtained from Adaçal Industrial Minerals Company (Afyon, TURKEY). Glycerol was also purchased from Merck. Citric acid was obtained from Aksu Company, TURKEY. Tween 80 was obtained from Sigma-Aldrich Company. The solutions were prepared using distilled water.

### B. Preparation of the biocomposite films

The preparation of three different biocomposite films was carried out using the solution casting method (Figure 1) according to the method of Terzioglu and Sıcak (2021) with some modifications [15]. The PVA powder (8% w/v) and

distilled water were mixed using a hotplate magnetic stirrer at 80°C. Cassava starch (2% w/v) was gelatinized in distilled water. The two mixture was mixed and then stirred at 600 rpm and heated at 70 °C for 60 min. After the mixture cooled to 50 °C, citric acid solution (10 % wt. of total polymer weight) was added and stirred for 30 minutes. Furthermore, glycerol (20%, wt. of total polymer weight) and Tween 80 was poured into the mixture with additional constant stirring for 15 minutes, respectively. At last, nano-CaCO<sub>3</sub> powder at concentrations of 1 and 2 % (wt. of total polymer weight) was added to the PVA/cassava starch film-forming mixture and stirred for 30 minutes. The 35 mL of obtained film forming mixture was poured into a 12 cm glass petri dish and dried at 40 °C for 24 hours. PVA/cassava starch without calcium carbonate was also prepared as a control film. The films prepared with 0, 1 and 2 % nano-CaCO<sub>3</sub> were named PSC-0, PSC-1 and PSC-2, respectively.

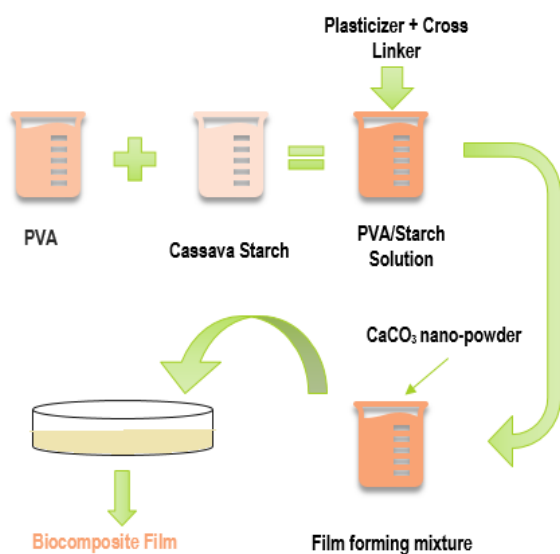


Figure I. Schematic diagram of PVA/cassava starch based biocomposite film preparation.

### C. Characterization of the biocomposite films

**FTIR analysis:** FTIR is a frequently used method to analyze functional groups in materials such as biocomposite films. FTIR spectra of biocomposite films were recorded with an FTIR spectrometer (Thermo Nicolet iS50, USA) within the wavenumber range of 4000-500 cm<sup>-1</sup> under 4 cm<sup>-1</sup> resolution.

**Mechanical analysis:** The mechanical features of samples were determined following the ASTM D 882 procedure using a Shimadzu AGS-X tester. Mechanical tests were repeated five times for each sample.

**Ultraviolet-Visible-Near infrared spectroscopy analysis:** The UV-VIS-NIR spectral analysis was characterized using UV-Vis spectrophotometer (Shimadzu UV-3600, Japan) with transmittance spectra of in the range of 200-800 nm. The film samples were cut into 2.5 x 3.0 cm rectangles and put in

the spectrophotometer. The transparency of the biocomposite films were calculated according to the following equation:

$$\text{Transparency} = \frac{\log(\%T_{600})}{y} \quad (1)$$

where y is the thickness of film (mm) and %T<sub>600</sub> is the percent transmittance at 600 nm [16].

## III. Results and Discussion

### A. Structural Properties of the biocomposite films

FTIR spectroscopy was used to determine the interaction between the PVA, cassava starch and nano-CaCO<sub>3</sub>. The FTIR spectrum of biocomposite films is presented in Figure II.

The spectrum of PSC-0 displayed a characteristic band at 1712 cm<sup>-1</sup> attributed to the C=C stretching vibration, which is typical for the ester bond and carboxyl groups in citric acid [17]. In the infrared spectra of films, stretching vibrations of O-H groups were at 3285 cm<sup>-1</sup> [18]. The peaks located at 1420, 1373 and 1324 cm<sup>-1</sup> were related to the CH<sub>2</sub> bending, CH<sub>2</sub> deformation and CH<sub>2</sub> stretching vibrations, respectively [2],[19]. The peak at 1240 cm<sup>-1</sup> representing the C-H wagging vibrations [18], while the peak at 1085 cm<sup>-1</sup> can be assigned to C-O stretching in C-O-H [18]. The band occurred at 841 cm<sup>-1</sup> due to the rocking vibration of CH<sub>2</sub> and the asymmetric stretching of C-O-C [20].

The pattern of nano-CaCO<sub>3</sub> incorporated samples (PSC-1 and PSC-2) showed small differences. The peaks at 3281 and 1420 cm<sup>-1</sup> were slightly shifted and the intensities of these peaks were lowered when compared to PSC-0. The peaks at ~ 1483, 876, and 712 cm<sup>-1</sup> are assigned to CaCO<sub>3</sub> (Figure not given) [21]. Therefore, the new peak formation that occurred at ~1580 cm<sup>-1</sup> could be related to the interaction of nano-CaCO<sub>3</sub> and the polymer matrix.

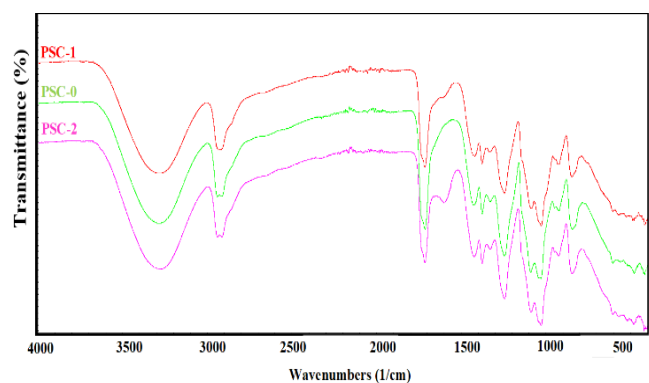


Figure II. FTIR spectrum of biocomposite films.

### B. Mechanical Properties of the biocomposite films

The mechanical properties of the films are given in Figure III. The calculated values of tensile strength, elongation at break, and Young's modulus of the biocomposite films are summarized in Table I. Tensile strength and elongation at break values varied in the range of 22.50±1.7-27.74±1.1 MPa and 186.75±17.3%-278.08±14.4%, respectively. The data showed that 1% nano-CaCO<sub>3</sub> incorporation increased the elongation capacity and flexibility of the control films as well as not changed the tensile strength. However, the further



increase of nanofiller content from 1% to 2% caused a decrease in the elongation at break and tensile strength values. This should be due to the agglomeration of the nanofiller at higher concentrations [16]. Young's modulus of the films was also varied in the range of  $38.28 \pm 2.0$ - $82.30 \pm 5.3$  MPa. Specifically, the nano- $\text{CaCO}_3$  at 1% led to a drop in Young's modulus value, while an increment occurred as its content reached 2%.

The capability to deform is particularly favorable in some industries including agriculture, cosmetics, and food packaging to fabricate elastic and flexible products based on the applications [22]. Therefore, when the elongation at break results were examined, the 1% nano- $\text{CaCO}_3$  loaded sample should be suggested as the best flexible film among the produced three samples.

Table I. Mechanical properties of biocomposite films.

Sample	Tensile strength (MPa)	Elongation at break (%)	Young's modulus (MPa)
PSC-0	$27.74 \pm 1.1$	$258.10 \pm 1.9$	$61.08 \pm 1.6$
PSC-1	$27.54 \pm 1.9$	$278.08 \pm 14.4$	$38.28 \pm 2.0$
PSC-2	$22.50 \pm 1.7$	$186.75 \pm 17.3$	$82.30 \pm 5.3$

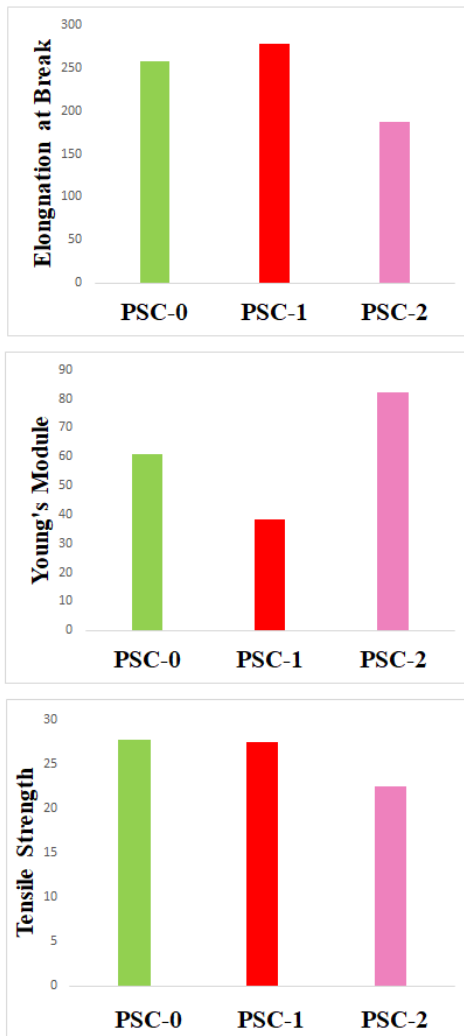


Figure III. Tensile strength, elongation at break and Young's modulus graphs of biocomposite films.

### C. Optical Properties of the biocomposite films

The transparency capacity of the films is a significant quality parameter. Because foods tend to oxidation at 200–280 nm that causes oxidative deterioration, discoloration, and off-flavor or rancidity [16]. It is expected from food packaging films not only to have UV-proof properties but also be transparent enough for visual lights [16],[23]. Therefore, determining the transmission percentage and transparency of films is important.

The visual appearance of biocomposite films was presented in Figure IV. The transparency of the films was found to be 6.65, 7.36 and 7.47 for PSC-0, PSC-1 and PSC-2, respectively. The nanofiller incorporation slightly increased the transparency of the films (Figure V). This result may be related to the reduction of the matrix crystallinity [24]. The transparency of PVA/cassava starch film (PSC-0) was slightly lower than the transparency of PVA/corn starch film (6.71) developed by Terzioğlu and Parin [17]. The addition of nano- $\text{CaCO}_3$  to the PVA/cassava starch matrix resulted in a small change in the transmittance in the range of 200-300 nm. Additionally, the transmission rate of the biocomposite film notably increased in the visible area, especially with 2% nano- $\text{CaCO}_3$  addition. It was shown that considerable transparency was still obtained (89~90% at 800 nm) for all developed PVA/starch biocomposite films.

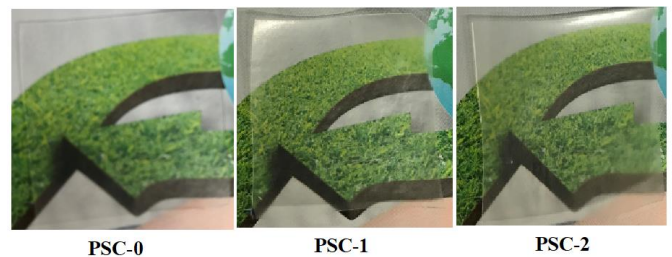


Figure IV. Visual appearance of biocomposite films.

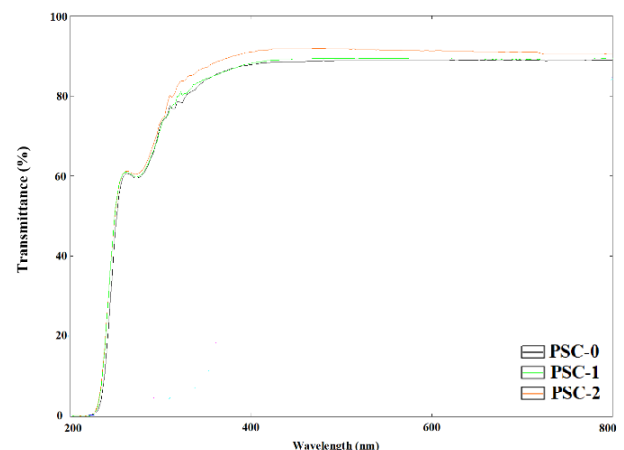


Figure V. UV-VIS-NIR spectra profiles of biocomposite films.

### Acknowledgment

This research is financially supported by the Scientific and Technological Research Council of Turkey (TÜBİTAK) with the 2209-B Undergraduate Research Projects Grant. The authors acknowledge Adaçal Industrial Minerals Company for supplying the calcium carbonate.

#### IV. Conclusion

Nano-CaCO<sub>3</sub> filler incorporated PVA/cassava starch films were prepared successfully using a green route (solution casting method). FTIR characterization results demonstrated similar structural composition of all developed films with the same functional groups. This fact reflects the physical interactions of the nanofiller and the polymer matrix. The nanofiller incorporation at 1% enhanced the elongation at the break of the films. The transparency of the films increased with nanofiller incorporation at both concentrations. Further scientific research is still needed to find the best nano-CaCO<sub>3</sub> incorporation content for this polymer matrix to obtain nanobiocomposite films with better functional properties.

#### References

- [1] Natwat Srikhao, Pornnapa Kasemsiri, Artjima Ounkaew, Narubeth Lorwanishpaisarn, Manunya Okhawilai, Uraiwan Pongsa, Salim Hiziroglu, Prinya Chindaprasirt, "Bioactive Nanocomposite Film Based on Cassava Starch/Polyvinyl Alcohol Containing Green Synthesized Silver Nanoparticles", *Journal of Polymers and the Environment*, Volume 29, 2021.
- [2] Hairul Abral, Angga Hartono, Fadli Hafizulhaq, Dian Handayani, Ani Sugiarti, Obert Pradipta, "Characterization of PVA/cassava starch biocomposites fabricated with and without sonication using bacterial cellulose fiber loadings", *Carbohydrate Polymers*, Volume 206, 2019.
- [3] Farah Fahma, Sugiarto, Titi Candra Sunarti, Sabrina Manora Indriyani, and Nurmalisa Lisdiana, "Thermoplastic Cassava Starch-PVA Composite Films with Cellulose Nanofibers from Oil Palm Empty Fruit Bunches as Reinforcement Agent", *International Journal of Polymer Science*, Volume 2017, 2017.
- [4] Maria-Cristina Popescu, Bianca-Ioana Dogaru, Mirela Goanta, Daniel Timpu, "Structural and morphological evaluation of CNC reinforced PVA/Starch biodegradable films", *International Journal of Biological Macromolecules*, Volume 116, 2018.
- [5] M. Lubis, A. Gana, S. Maysarah, M.H.S. Ginting, M.B. Harahap, "Production of bioplastic from jackfruit seed starch (*Artocarpus heterophyllus*) reinforced with microcrystalline cellulose from cocoa pod husk (*Theobroma cacao* L.) using glycerol as plasticizer", *IOP Conference Series: Materials Science and Engineering*, 309, 2018.
- [6] Nataliya E. Kochkina, Olga A. Butikova, "Effect of fibrous TiO<sub>2</sub> filler on the structural, mechanical, barrier and optical characteristics of biodegradable maize starch/PVA composite films", *International Journal of Biological Macromolecules*, Volume 139, 2019.
- [7] Frédéric Chivrac, Eric Pollet, Patrice Dole, Luc Avérous, "Starch-based nano-biocomposites: Plasticizer impact on the montmorillonite exfoliation process", *Carbohydrate Polymers*, Volume 79, Issue 4, 2010.
- [8] HS Mahadevaswamy, B Suresha, Role of nano-CaCO<sub>3</sub> on mechanical and thermal characteristics of pineapple fibre reinforced epoxy composites, *Materials Today: Proceedings*, Volume 22, Issue 3, 2020.
- [9] Pınar Terzioğlu, "Electrospun Chitosan/Gelatin/Nano-CaCO<sub>3</sub> Hybrid Nanofibers for Potential Tissue Engineering Applications", *Journal of Natural Fibers*, 2021. DOI: 10.1080/15440478.2020.1870639
- [10] Kalyani Prusty, Sarat K Swain, "Nano CaCO<sub>3</sub> imprinted starch hybrid polyethylhexylacrylate/polyvinylalcohol nanocomposite thin films", *Carbohydrate Polymers*, Volume 139, 2016.
- [11] Carla Vilela, Carmen S.R. Freire, Paula A.A.P. Marques, Tito Trindade, Carlos Pascoal Neto, Pedro Fardim, "Synthesis and characterization of new CaCO<sub>3</sub>/cellulose nanocomposites prepared by controlled hydrolysis of dimethylcarbonate", *Carbohydrate Polymers*, Volume 79, Issue 4, 2010.
- [12] Sudhir K. Kisku, Niladri Sarkar, Satyabrata Dash, Sarat K. Swain, "Preparation of Starch/PVA/CaCO<sub>3</sub> Nanobiocomposite Films: Study of Fire Retardant, Thermal Resistant, Gas Barrier and Biodegradable Properties", *Polymer-Plastics Technology and Engineering*, Volume 53, Issue 16, 2014.
- [13] Norio Fukuda, Hideto Tsuji, Yasushi Ohnishi, "Physical properties and enzymatic hydrolysis of poly(L-lactide)-CaCO<sub>3</sub> composites", *Polymer Degradation and Stability*, Volume 78, 2002.
- [14] Qingjie Sun, Tingting Xi, Ying Li, Liu Xiong, "Characterization of Corn Starch Films Reinforced with CaCO<sub>3</sub> Nanoparticles", *Plos One*, Volume 9, Issue 9, 2014.
- [15] Pınar Terzioğlu, Yusuf Sicak, "Citrus Limon L. Peel Powder Incorporated Polyvinyl Alcohol/Corn Starch Antioxidant Active Films", *Journal of the Institute of Science and Technology*, Volume 11, Issue 2, 2021.
- [16] Shaoxiang Lee, Meng Zhang, Guohui Wang, Wenqiao Meng, Xin Zhang, Dong Wang, Yue Zhou, Zhonghua Wang, "Characterization of polyvinyl alcohol/starch composite films incorporated with p-coumaric acid modified chitosan and chitosan nanoparticles: A comparative study", *Carbohydrate Polymers*, Volume 262, 2021.
- [17] Pınar Terzioğlu, Fatma Nur Parm, "Polyvinyl Alcohol-Corn Starch-Lemon Peel Biocomposite Films as Potential Food Packaging", *Celal Bayar University Journal of Science*, Volume 16, Issue 4, 2021.
- [18] Phetdaphat Boonsuk, Apinya Sukolrat, Kaewta Kaewtatip, Sirinya Chantarak, Antonios Kelarakis, Chiraphon Chaibundit, "Modified cassava starch/poly(vinyl alcohol) blend films plasticized by glycerol: Structure and properties", Volume 137, Issue 26, 2020.
- [19] Priyanka Rani, M Basheer Ahamed, Kalim Deshmukh, "Dielectric and electromagnetic interference shielding properties of carbon black nanoparticles reinforced PVA/PEG blend nanocomposite films", *Materials Research Express*, Volume 7, 2020.
- [20] Anida M.M. Gomes, Paloma L. da Silva, Carolina de L. e Moura, Claudio E.M. da Silva, Nágila M.P.S. Ricardo, "Study of the Mechanical and Biodegradable Properties of Cassava Starch/Chitosan/PVA Blends", *Macromolecular Symposia*, Volume 299-300, Issue 1, 2011.
- [21] S. El-Sherbiny, S.M. El-Sheikh, A. Barhoum, "Preparation and modification of nano calcium carbonate filler from waste marble dust and commercial limestone for papermaking wet end application", *Powder Technology*, Volume 279, 2015.
- [22] H.P.S. Abdul Khalil, E.W.N. Chong, F.A.T. Owolabi, M. Asniza, Y.Y. Tye, H.A. Tajarudin, M.T. Paridah, S. Rizal, Microbial-induced CaCO<sub>3</sub> filled seaweed-based film for green plasticulture application, *Journal of Cleaner Production*, Volume 199, 2018.
- [23] Azam Akhavan, Farah Khoylou, Ebrahim Ataeivarjovi, "Preparation and characterization of gamma irradiated Starch/PVA/ZnO nanocomposite films", *Radiation Physics and Chemistry*, Volume 138, 2017.
- [24] Vincenzo Titone, Francesco Paolo La Mantia, Maria Chiara Mistretta, "The Effect of Calcium Carbonate on the Photo-Oxidative Behavior of Poly(butylene adipate-co-terephthalate)", *Macromolecular Materials and Engineering*, Volume 305, Issue 10, 2020.

# Movie Reviews Text Sentiment Analysis Based On Hybrid LSTM and GloVe

Nour Ammar  
Department of Software Engineering  
Istanbul Aydin University  
Istanbul, Turkey  
noor101ammar@gmail.com

Ali Okatan  
Department of Software Engineering  
Istanbul Aydin University  
Istanbul, Turkey  
aliokatan@aydin.edu.tr

**Abstract**— *Social media has come to be a useful resource for critiques and ratings. The mining of vital information that allows governments to keep public protection and organization's growth income has gradually advanced. However, the performance of current techniques requires continuous improvement due to the exponential growth of information. In this paper, the proposed structure is the LSTM neural network, an advanced spatial technique of RNN. The verbal language information has been utilized by changing it into numerical information relying upon the GloVe dictionary. The manner utilized in natural language processing has accomplished acceptable results, and the reasonable prediction changed into real with an excessive speed.*

**Keywords**—NLP, LSTM, RNN, GloVe

## I. Introduction

Recently, blogs, websites, and social media are considered a powerful means of collecting product reviews. The proposed method achieves maximum benefit at minimum cost in minimum time by analyzing users' opinions on positive and negative reviews using Long Short Term Memory (LSTM) algorithms. The required data can be extracted by technical methods as in the proposed method.

Recurrent Neural Networks (RNNs) introduced by [14] provides high performance in speech processing, speech recognition, speech translation, stock prediction and semantic analysis, used in this study. LSTM neural network is a modified architecture of traditional RNN and is more accurate [1]. The change from RNN to LSTM is the constant backpropagated error flow in gradients processed in the LSTM network [13]. An LSTM layer consists of a set of recurrently connected blocks called memory blocks. LSTM can learn to bridge minimal time delays of more than 1000 discrete time steps by enforcing a constant error flow through Constant Error Carousels (CECs) [13]. In the presented approach, a model consisting of two LSTM layers with a total of seven layers is used. The results were fairly high and satisfying; accuracy is high of each training and testing as well as the speed of testing as detailed in section IV.

## II. Related Work

In [1], a text sentiment analysis based on LSTM model was presented to analyse human emotions. The training data is classified into three categories (negative, positive, and neutral) according to emotions, and then fitted into the LSTM models trained for each data category, resulting in multiple LSTM models for the corresponding emotional ratings. The accuracy in [1] was higher than traditional RNNs. Another approach is followed in [2], where both models are used, Convolutional Neural Network (CNN) with RNN-LSTM. Using the CNN model is to extract local features. The goal of using the LSTM model is to capture long-distance dependencies and combine the extracted features into a single hybrid CNN-LSTM model. In [2], obtained efficient results. Each memory block contains at least one recurrently

connected memory cell and three gates, namely the input, output, and forget gates. The gates are multiplicative units and enable similar continuous write, read and reset operations. The network can only interact with the cells through the gates [2].

In [3], a solution approach to the weakness of LSTM networks processing continuous input streams was proposed by the "learn to forget" algorithm, which refers to a novel adaptive "forget gate". The "forget gate" allows an LSTM cell to learn to reset itself at appropriate times. The approach proposed in this study is very similar to the approach of [2].

All the above approaches refer to [6], the approach presented in 1997 by Hochreiter & Schmidhuber, who were the first to define LSTM as a special kind of RNN with the ability to learn long-term dependencies.

## III. Overview, Methods and Tools

The RNN-LSTM model was trained on the IMDB movie reviews dataset as a function of the GloVe word embedding dictionary.

### A. GloVe words embedding dictionary

GloVe [5] is a word vectorization technique that embeds words into a concise vector space where similar words are found close to each other as clusters, while different words are found far away from each other, as shown in Figure I. GloVe embedding is preferable to Word2vec embedding because GloVe relies on local statistics (local context information of words) while it incorporates global statistics (word co-occurrence) to obtain word vectors. The dimension of the data is 200. Figure I shows the 3D graph of the embedded words after conversion to sequences by PCA using the tensorboard tool provided by Tensorflow.

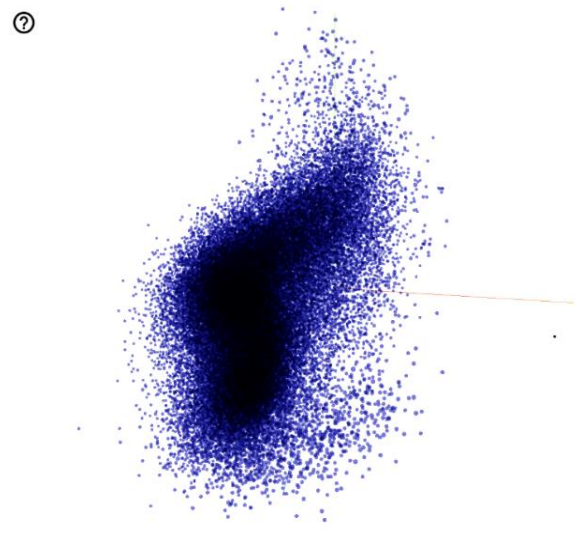


Figure I. 3-D diagram PCA by tensorboard of GloVe embeddings sequences

## B. RNN-LSTM architecture

Recurrent neural networks (RNNs) [14] are related to sequences, lists, and continuously initialized data. There have been incredible successes in applying RNNs to a variety of problems: Speech recognition, language modeling, translation, image labeling, and more. LSTMs - introduced in [6] by Hochreiter & Schmidhuber, are a special form of RNNs that avoids the problem of long term dependency that exists in traditional RNNs, it utilizes tanh activation (1). Moreover, the LSTM model contains four neural network layers as shown in Figure II, which makes the model more robust to deal with problems such as the vanishing gradient problem and reduce long dependencies by selecting only essential information from the previous cells.

$$T(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (1)$$

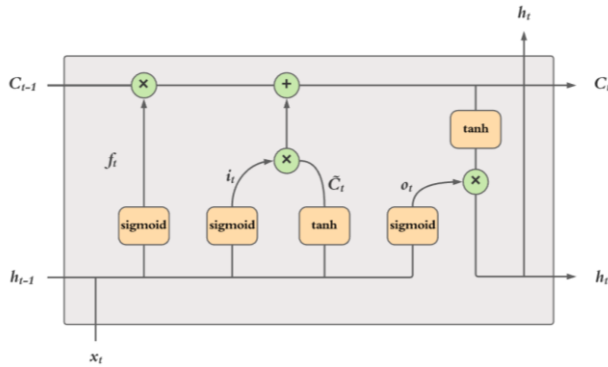


Figure II. LSTM block showing four interacting layers, gates, and operations

Figure II clarifies the process of LSTM block's gates-input, forget, output, and updated Cell-and the parameters-activation functions-of gates as shown in (2).

$$\begin{pmatrix} i \\ f \\ o \\ \tilde{C}_t \end{pmatrix} = \begin{pmatrix} \sigma \\ \sigma \\ \sigma \\ \tanh \end{pmatrix} W \begin{pmatrix} h_{t-1} \\ x_t \end{pmatrix} \quad (2)$$

The steps of the operation flow in LSTM block start with Forgetting irrelevant past information using (3). Then, identify New Information to be Stored, the sigmoid layer, which indicated in (4), decides what values to update, tanh layer which indicated in (5) generate a new vector of values that could be added to the state called "candidate values". Then apply forget operation to the previous internal cell state by (6) And add new candidate values, scaled by the number of values that decided to be updated by (7). After that, the two operations are summited to produce (8) that updates the cell state. After updating the state, last sigmoid layer (9) decides what parts of state to output. Finally, output a filtered cell state (10) that uses tanh layer (1) to squash values between 1 and -1.  $ht$  is in the same time an input to the next cell in the continuous LSTM block chain [6, 7, 16, 17].

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f) \quad (3)$$

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i) \quad (4)$$

$$\tilde{C}_t = \tanh(W_c[h_{t-1}, x_t] + b_c) \quad (5)$$

$$f_t \times C_{t-1} \quad (6)$$

$$i_t \times \tilde{C}_t \quad (7)$$

$$C_t = f_t \times C_{t-1} + i_t \times \tilde{C}_t \quad (8)$$

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \quad (9)$$

$$h_t = o_t \times \tanh(C_t) \quad (10)$$

The project was implemented using the TensorFlow platform in python; it is a high-performance computations platform.

## C. Deep Neural Network Architecture

Figure II, shows the basic network structure; however, the final architecture is given in Figure III. The proposed model contains seven layers, which include both the input and the output. To avoid overfitting and improve performance, a dropout process is used. The input and recurrent connections to the LSTM units are excluded from the activation and weight updates during network training [8]. The dense layer is used as an output layer with an activation function to improve the performance of the approach [9].

### 1) Layers:

1. A sequence of words with 256x1 dimension is used as an input to the network.
2. The input of the second layer is the matrix of embeddings multiplied by sequences. The output dimension is 256x200.
3. Here is the first LSTM layer with 128 features (neurons). The output dimension is 256x128.
4. First dropout layer, input dimension and output dimension are the same but it is used to reduce overfitting.
5. Second LSTM layer with 32 features with dimension 32x1. 32 are the number of embedded sequences reduced from 200 gradations to 32 gradations.
6. The second dropout layer, input, and output dimensions are the same, and it is used to reduce overfitting.
7. The output layer is a Dense layer with dimension 1x1 as positive or negative weighting.

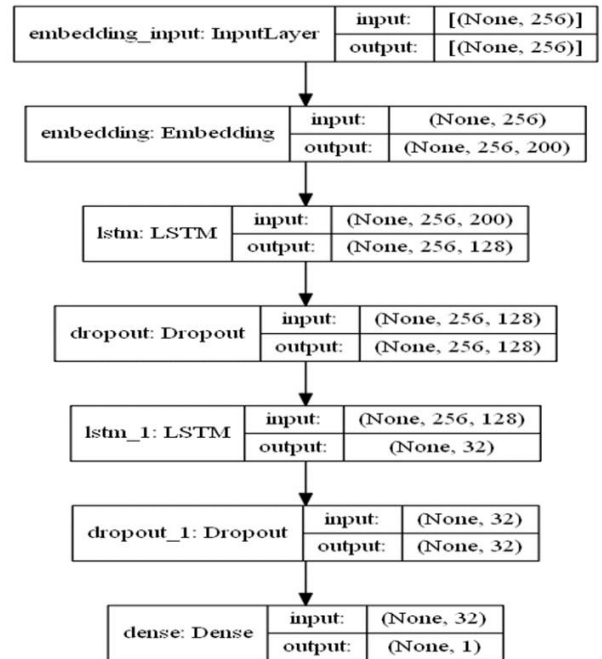


Figure III. Architecture of neural network layers showing number of input and output in each layer

2) *Functions:*

a) *Activation function:*

There are many activation functions, which can be used with the proposed model; however, the sigmoid activation function is chosen in this case. Figure IV shows the sigmoid function. The sigmoid function is represented by (11).

$$\sigma = \frac{1}{1 + e^{-x}} \quad (11)$$

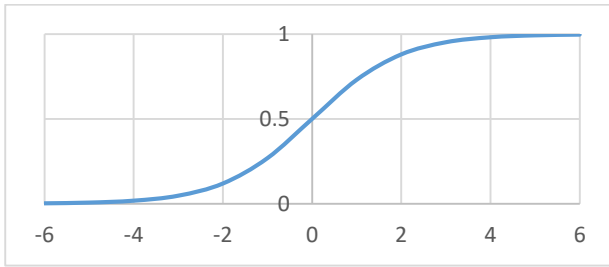


Figure IV. Sigmoid function curve that lies between 0 and 1

b) *Optimization function:*

The optimization function used in this study is Adam optimization, introduced by [11]. It is an adaptive learning rate-based method, which means it computes individual learning rates for different epochs. Adam refers to adaptive moment estimation. It uses estimates of the first and second moments of the gradient to adapt the learning rate for each weight of the neural network. The Nth moment is a random variable used as the expected value of that variable to the power of n.

$$\theta_{t+1} = \theta_t - \frac{\eta}{\sqrt{\hat{v}_t + \epsilon}} \hat{m}_t \quad (12)$$

c) *Loss function:*

The mean squared error (MSE) is used as a fitness function to minimize the error. It calculates the mean of the squared differences between the predicted and actual values. It minimizes the loss value [12].

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (13)$$

3) *Data:*

IMDB is a popular movie website. It combines movie plot description, metastore ratings, critic and user ratings and reviews, release dates, and many other aspects. The system is trained using the IMDB dataset. This is a CSV file format that contains 50K textual movie reviews with high polarity for semantic analysis and Natural Language Processing (NLP). The data fields are the review and its value, which is either 0 or 1, where 0 is negative and 1 is positive, as shown in Table I. This dataset is freely available [4]. The data is preprocessed and tokenized using the libraries nltk and textblob.

Table I. sample of IMDB dataset

	review	sentiment
0	One of the other reviewers has mentioned that ...	positive
1	A wonderful little production.   The...	positive
2	I thought this was a wonderful way to spend ti...	positive
3	Basically there's a family where a little boy ...	negative
4	Petter Mattei's "Love in the Time of Money" is...	positive
5	Probably my all-time favorite movie, a story o...	positive
6	I sure would like to see a resurrection of a u...	positive
7	This show was an amazing, fresh & innovative i...	negative
8	Encouraged by the positive comments about this...	negative
9	If you like original gut wrenching laughter yo...	positive
10	Phil the Alien is one of those quirky films wh...	negative

4) *Training and Testing*

In this study, the dataset was divided into three parts: 70% for training, 20% for testing, and 10% for validation. The network is trained depending on the GloVe embedding dictionary [5]. The training time is machine hardware dependent. The training time and performance of the model were improved by applying five basic measures, including:

- Adding and discarding layers of the model as replacing the output layer with a dense layer.
- Using the TensorFlow platform, which is suitable for high performance computations.
- Use of sigmoid activation function.
- Use of Adam optimization function.
- Use of MSE loss function.

**IV. Experiments, Analysis and Performance**

The model was implemented using the open source framework TensorFlow. Training was performed on a PC with SSD hard drive 16 RAM and i7 10500 core.

**A. Platform**

Training the network took about 20 minutes using the TensorFlow environment. The flow of tensors (computations of the model) can be viewed using the Tensorboard tool provided by TensorFlow, as shown in Figure V Each container (tensor) is a constant, vector or higher dimensional data and each arrow (flow) is a directed flow of computations.

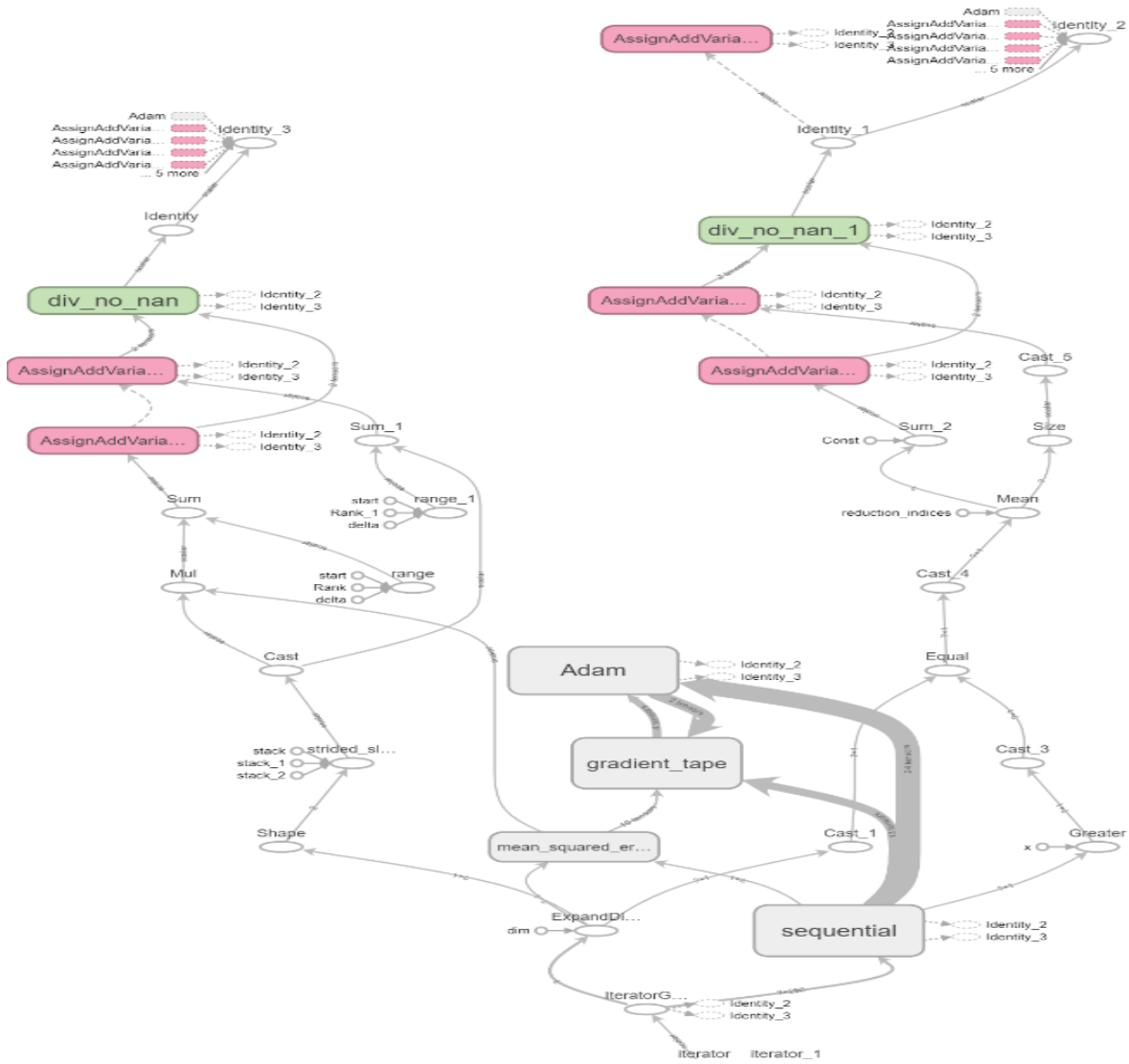


Figure V. Tensorboard graph showing computations flow of the model

### B. Results

Table II shows the values of the confusion matrix of the model training, which, as can be seen, the true values (blue) are higher than the negative values (red). Table III shows the ratios of the training, testing, and validation results. The accuracy of the model is 87.1%.

Figure VI and Figure VII show the plots for accuracy and loss during the training and testing process. As shown, the curves of training and testing increase in the accuracy plot, while in the loss plot, the curves decrease, which confirms the excellent performance of the model.

Table III. Values of training, testing, and validation ratios

	training	testing	validation
<b>Accuracy</b>	86.24%	87.12%	87.19%
<b>Loss</b>	9.88%	11.25%	10.09%
<b>Score</b>	-	10.30%	10.21%
<b>F1 score</b>	-	86.32%	-

Table II. Confusion matrix of the trained model

	positive	negative
positive	4647	942
negative	346	4065

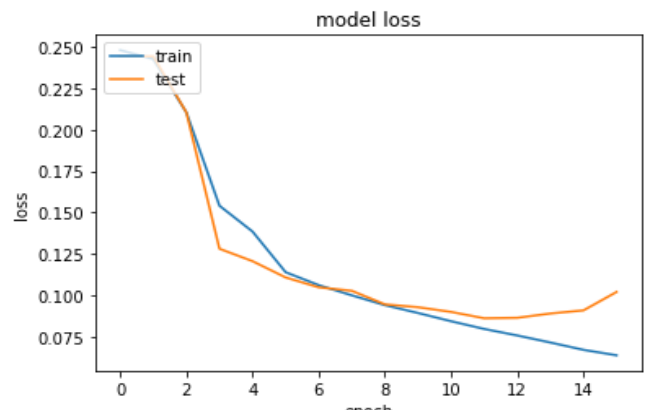


Figure VI. Accuracy of model during training and testing in 14 epochs

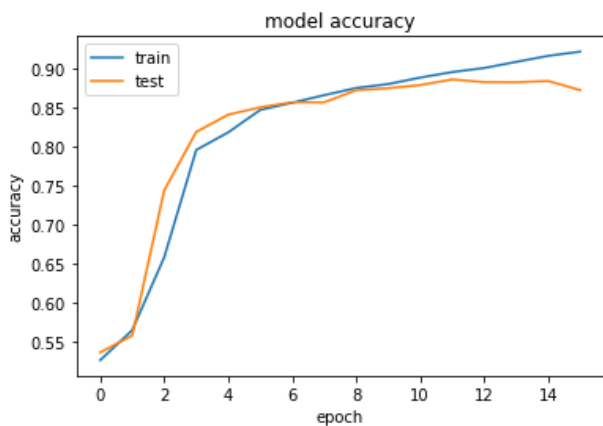


Figure VII. Loss of model during training and testing in 14 epochs

```

x=[ 10 45 795 23 19 332 1 261 455 105 30 4844 327 1
1292 213 201 88 683 35 67 214 48 1 1177 2 35 67
1010 4 71 6 198 10 115 3953 2 1 698 952 51 4
1 453 4302 67 244 3210 78 188 35 67 44 540 381 32
44 3 1 795 441 187 435 81 10 731 6 63 49 721
1042 395 4 1 502 3912 1542 2 6 61 264 1 854 43
13 1 445 353 159 2 445 1465 1 4227 381 67 652 964
8 10 1612 29 24 3 1 84 2392 204 4509 7 1292 266
19 11 1152 11 1 24 7 24 102 55 57 493 64 3998
1 107 5 1282 13 26 1 3113 4080 150 19 11 4469 11
55 2 7 44 91 9 5 45 146 1 109 10 49 45
597 5 116 46 351 8 643 1392 15 46 98 3 123 643
4 1643 22 237 7 1 203 11 2 1588 2 75 1351 4
62 97 2203 2 1 1423 63 77 3326 2 65 1399 18 1
218 559 1847 11 4413 26 7 26 1199 795 15 153 3825 335
1 114 4441 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0], Predicted=[0.96169996]
positive review

```

Figure VIII. example of a sample sequence in test dataset of predicting with probability of prediction

Figure VIII shows shows selected normalized data sample and the probability of prediction, which is a true prediction. Figure IX shows a practical example done by the user. In this case, the executed program prompts the user to input a sentence then the system preprocesses the input and predicts within milliseconds only.

```

Enter your value: i liked the movie so much. i will watch it again. i will invite my friends to watch it.

preprocess_text(TEST_REVIEW)
remove_tags(TEST_REVIEW)
TEST_REVIEW = tokenizer.texts_to_sequences(TEST_REVIEW)
TEST_REVIEW = pad_sequences(TEST_REVIEW, padding='post', maxlen=maxlen)

y_pred = lstm_model.predict(TEST_REVIEW)
if y_pred[0] < 0.5:
    value = 'negative review'
else:
    value = 'positive review'
print("Predicted=%s, %s" % ( y_pred[0], value ))

Predicted=[0.60894475], positive review

```

Figure IX. Practical example of testing the model

## V. Conclusion

The results show that the use of the LSTM improved performance of both speed and accuracy is improved. The GloVe 200-D dictionary was embedded in the Tensorflow environment. Tensorflow has been very successful in handling high-speed computations. It is expected that the model presented in this study can be modified to be used for real-time emotion in speech recognition.

## References

- [1] Alex Graves, Jürgen Schmidhuber, Framewise phoneme classification with bidirectional LSTM and other neural network architectures, *Neural Networks*, Volume 18, Issues 5–6, 2005, Pages 602-610, ISSN 0893-6080, <https://doi.org/10.1016/j.neunet.2005.06.042>.
- [2] Rehman, A.U., Malik, A.K., Raza, B. *et al.* A Hybrid CNN-LSTM Model for Improving Accuracy of Movie Reviews Sentiment Analysis. *Multimed Tools Appl* **78**, 26597–26613 (2019). <https://doi.org/10.1007/s11042-019-07788-7>
- [3] Felix A. Gers, Jürgen Schmidhuber, Fred Cummins; Learning to Forget: Continual Prediction with LSTM. *Neural Comput* 2000; 12 (10): 2451–2471. doi: <https://doi.org/10.1162/089976600300015015>.
- [4] Leon, Stefano, 2020, IMDb movies extensive dataset <https://www.kaggle.com/lakshmi25npathi/imdb-dataset-of-50k-movie-reviews>.
- [5] Pennington, Jeffrey & Socher, Richard & Manning, Christopher. (2014). Glove: Global Vectors for Word Representation. *EMNLP*. 14. 1532-1543. 10.3115/v1/D14-1162.
- [6] Hochreiter, Sepp & Schmidhuber, Jürgen. (1997). Long Short-term Memory. *Neural computation*. 9. 1735-80. 10.1162/neco.1997.9.8.1735.
- [7] Karpathy, A., Johnson, J., & Fei-Fei, L. (2015). Visualizing and understanding recurrent networks. *arXiv preprint* <https://arxiv.org/abs/1506.02078>.
- [8] Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: a simple way to prevent neural networks from overfitting. *The journal of machine learning research*, 15(1), 1929-1958.
- [9] Applications of Deep Neural Networks, Jeff Heaton Updated regularly: last update: 21 Jan 2021 (v2) <https://arxiv.org/abs/2009.05673>.
- [10] Han J., Moraga C. (1995) The influence of the sigmoid function parameters on the speed of backpropagation learning. *Lecture Notes in Computer Science*, vol 930. Springer, Berlin, Heidelberg. [https://doi.org/10.1007/3-540-59497-3\\_175](https://doi.org/10.1007/3-540-59497-3_175).
- [11] Kingma, D. P., Ba, J. (2014). Adam: A method for stochastic optimization. <https://arxiv.org/abs/1412.6980>.
- [12] Alejo, R., García, V. & Pacheco-Sánchez, J.H. An Efficient Over-sampling Approach Based on Mean Square Error Back-propagation for Dealing with the Multi-class Imbalance Problem. *Neural Process Lett* 42, 603–617 (2015). <https://doi.org/10.1007/s11063-014-9376-3>.
- [13] Ravi, Vinayakumar & Kp, Soman & Poornachandran, Prabakaran. (2017). Evaluation of Recurrent Neural Network and its Variants for Intrusion Detection System (IDS). *International Journal of Information System Modeling and Design*. 8. 43-63. <https://doi.org/10.4018/IJISMD.2017070103>.
- [14] Rumelhart, David E., Geoffrey E. Hinton, and Ronald J. Williams. *Learning internal representations by error propagation*. California Univ San Diego La Jolla Inst for Cognitive Science, 1985.
- [15] Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. <https://arxiv.org/abs/1301.3781>.
- [16] Li, F., Johnson, J., Yeung, S. (2017). Lecture 10: Recurrent Neural Networks [PowerPoint slides]. Retrieved from Stanford University School of Engineering. Convolutional Neural Networks for Visual Recognition. CS231 n. CS231n. [http://cs231n.stanford.edu/slides/2017/cs231n\\_2017\\_lecture10.pdf](http://cs231n.stanford.edu/slides/2017/cs231n_2017_lecture10.pdf).
- [17] Jiang, C., Chen, Y., Chen, S., Bo, Y., Li, W., Tian, W., & Guo, J. (2019). A Mixed Deep Recurrent Neural Network for MEMS Gyroscope Noise Suppressing. *Electronics*, 8(2), 181. MDPI AG. Retrieved from <http://dx.doi.org/10.3390/electronics8020181>.

# Simulation Comparative Study to Highlight the Relation Between Building Form and Energy Consumption

Omar ALGBURI  
 Department of Architecture Engineering  
 Istanbul Sabahattin Zaim University  
 Istanbul, Turkey  
 omar.algburi @izu.edu.tr

Bahar FERAH  
 Department of Architecture Engineering  
 Istanbul Sabahattin Zaim University  
 Istanbul, Turkey  
 baharak.fareghi@izu.edu.tr

**Abstract**— Buildings construction is one of the causes of climate change. Architects who design energy-efficient and sustainable buildings lead in establishing the solutions to climate change. Decisions concerning building form design and its impact on energy efficiency are the main motive behind this work. This study examines the role of building form on energy consumption on two proposed entire glazed office buildings, cubic (rectangular plan layout) and cylinder (circular plan layout). AutoCAD software was utilized to draw the plan layout, then the thermal simulation and the proposed two forms were modeled and analyzed respectively using DesignBuilder software. The results show that the energy consumption of an entire glazed façade building strongly depends on its solar heat gain through exterior windows, building orientation, and glazing shading amount. Furthermore, building form amongst the studied factors plays the most critical role in determining a building's energy consumption. In the end, suggestions were presented to raise the awareness of architects to design energy-efficient buildings form.

**Keywords**— Building form, Climate change, DesignBuilder, Energy simulation

## I. Introduction

Is there a significant and direct effect of the building Form on energy consumption considered the central question behind this study? In the first stage of designing a building (conception stage), engineers and architects are facing numerous economic, social, techniques details, ecological issues, and aesthetic limitations. Indeed, designing a building is a complicated task, and to be energy efficient is unique. Moreover, the concern for pollution reduction, climate change fighting, and energy savings must remain one of the principles designers should regards. When the designers develop their project's concept in the first stage, they need operative knowledge about energy-efficient design principles.

The form and size of a building have a sufficiently significant influence on energy consumption [1],[2]. [3]. As stated in the study of Kocagil et al. [4], the building Form and the envelope complexity directly impact the total heat loss and gain and consequently the energy consumption. The proportion between the building's thermal envelope area (A) and its volume (V) is known as the A/V ratio. The Form factor of a building is relying on this proportion which measures the compactness of the building. The thermal envelope represents the isolation part between the outdoor (the unconditioned area) and the indoor (the conditioned space) environment. The value of this Form factor determines the form of the building for a given volume. The study of Danielski et al. (2012) explained the value of the Form factor as illustrated in Figure I. As shown in this

figure, the unit of the Form length represented by the parameter 'a' symbolizes. Although all the buildings (A, B, C, D) have similar volumes, they have different Form factors because of their different thermal envelope areas [2]. The Form factor also relies on the building size. Building (C) is larger but has a lower form factor than building (A), although they have the same form. Irregular Forms with open balconies that extend beyond the façade may also increase the Form factor, as illustrated by buildings D in Figure I.

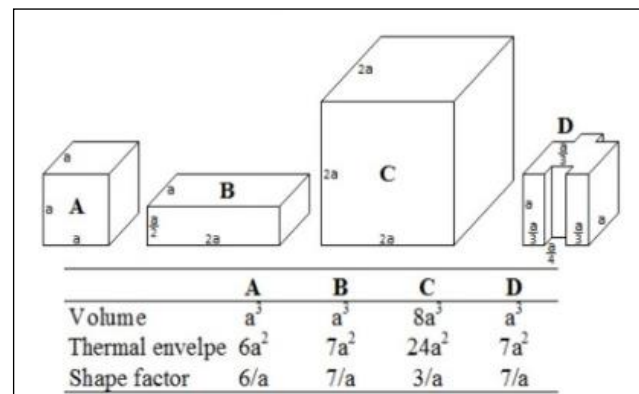


Figure I. Different sizes and Forms with the Form factor of each building [2].

The concept of building Form factor is linked to heat losses and gain on buildings, resulting in energy consumption[3]. Building with a larger envelop surface area in proportion to their volume will have a higher Form factor, maximizing heat losses. In contrast, buildings with lower Form factors need lower specific heat demand[5].

Many studies addressed the building geometry and Form factors that affect the energy consumption of buildings[6][7][8]. The study done by [8] proposed a methodology to determine building form related to opaque component U-value represented by A/V ratio. A survey conducted by Joelsson et al. (2012), it was shown the vital impact of the Form factor on final energy demand in residential buildings [2]. Another study [9] shows that space boundaries and building geometry significantly influence energy consumption factors. In addition, in the study done by Al-Anzi [6] et al., different office building Forms have been developed to study each form's energy demand. Many studies were done using the energy simulation method for optimizing the building Form in terms of its energy consumption [10][11][12][13][14][15][16]. However, many factors can impact the energy consumption of a building. For example, according to (ASHRAE 90.1) American Society of Heating, Refrigerating, and Air-Conditioning



Engineers [17] determined five main factors that impact energy consumption of a building as follows: (1) mechanical systems ; (2) building envelope ; (3) power generation systems ; (4) water heating and (5) lighting systems. The role of the designer here is to manage these factors and design energy-efficient buildings. Other factors such as window-to-wall ratio [18], glazing type [19] [20], solar heat gain coefficient (SHGC) [21] [22], thermal insulation of materials[23] [24], sun shading [25], and surface coloring [26] have a significant effect on the energy consumption of buildings.

### A. Energy consumption on buildings

To understand energy consumption in buildings, we should first understand how a building uses energy. According to many studies, most of the energy consumed in buildings is essentially used to enable an acceptable level of users' thermal comfort. Cooling or heating the indoor air using air conditioning units or providing fresh air by ventilation. Other energy uses are domestic hot water, artificial light, household appliances, or other electrical equipment (refrigerators, computers, TVs, etc.) Lighting, an essential part of every indoor environment, often consumes less energy than other electric appliances.

Heating or cooling anything - air, water, or food - is an energy-intensive process, and as such, appliances related to heating or cooling are often energy-intensive. Air conditioners and space heating appliances have an essential power requirement, as do cooking ovens, toasters, and microwaves. However, those devices like the blender, steam iron, and toaster are only used for a small time interval. That's leading to reduces energy used compared to appliances like the air conditioner, which is used for a much longer duration. However, the most extensive domestic energy requirement on the building is space heating and cooling. According to [24], wisely choosing the wall and the window material and the architecture of the enclosed space can significantly reduce the energy required for space heating and cooling. The heat energy from the outside air will be passed through the walls and windows to the inside of the building. The energy is transferred by vibrations of neighboring molecules in the wall and windows. This leads to an increase in indoor temperature. This mechanism of heat transfer is called conduction, denoted by Q. The power transferred through conduction is given by the following equation (1):

$$Q = U.A.\Delta T \quad (1)$$

Where,

Q: heat transferred in W

U: overall heat transfer coefficient in W/m<sup>2</sup>.°c

A: heat transfer area in m<sup>2</sup>

ΔT: temperature difference on each surface of the wall in °c

### B. Energy modeling and simulation

Energy modeling is the process of computerized the parameters of a building used to perform energy simulation. In contrast, energy simulation predicts building energy performance by using software analysis [27].

According to [27], building parameters such as orientation, thermal properties, and envelope properties contain computations factors that make energy simulation a

complex process. Energy simulation is mainly founded to help the designer make the right decisions in the early design phase. According to many recent studies evaluating and analyzing the building energy performance lead to design energy-efficient building. Therefore, design improvements or alternatives could be established by the design team to enables energy saving. However, making different building Form models in order to identify energy-saving alternatives does not typically occur correctly in many cases. In addition, adapting simulation models (i.e., input data) usually ends in various coding errors as it consists of manual or semi-manual representation data[28]. A gradual simulation method of a confirmed system with validation methodology of building energy analysis should be involved by an expert team[29].

## II. Research Methodology

The energy simulation method by DesginBuilder software was used to test the impact of the building form on the reduction rate of energy consumption. Energy simulation analyses were performed to evaluate the two proposed Cylinder (circular plan) and cubic (square plan) building forms in terms of energy consumption. Two-dimensional modelings of the buildings were the first step using Autocad. Importing the DXF model of the two different design alternatives using Designbuilder simulation software was the second step.

The simulation method is a sequence of three related phases, as shown in Figure II.

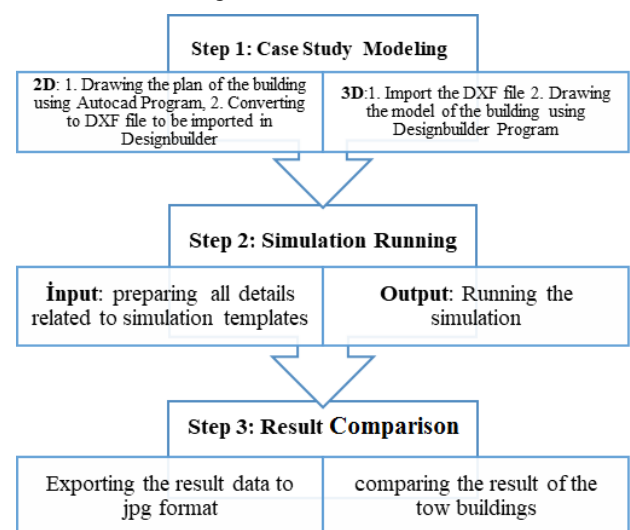


Figure II. A diagram shows the steps of the simulation methodology

This study used comparative thermal analysis between two different forms, cubic and cylinder office building, as shown in Figure III. For comparison needs, the two forms were proposed with the same floor area and same volume to evaluate the impact of the building form on its energy consumption. Both buildings have a central open courtyard to enable access to natural light and ventilation. The two proposed buildings are located in Istanbul- Turkey. The height was also the same on both, with Four typical floors 3.4 m in height for each floor. Therefore, the total height is considered to be 15 m, 3.5 \*4 (number of floors). The cubic building plan dimensions are 64m\*64m, and its total ground floor area was 4096 m<sup>2</sup>. Approximately the same ground floor area applied in the Cylinder form (circular plan). The radius of the circle is 36 m. The area of a circle is pi times

the radius squared ( $A = \pi r^2$ ). Therefore, the total ground floor area was calculated as  $(36 \times 36 \times 3.14) = 4069 \text{ m}^2$ .

Before running the simulation, parameters of the energy simulation processes should be identified first. The comparative simulation focused on cooling energy consumption during a specific period, the peak cooling load from 21 June till 21 September, as a typical summer season. Indoor operative temperature and solar heat gain through exterior windows were selected to determine the total cooling energy consumption during the summer season. These parameters have an essential influence on the average indoor operative temperature, overall cooling load, and energy consumption. Internal heat gain through the wall and roof was ignored.

The construction templates of both buildings in this study used to be Turkiye-medium weight which is data for a range of ready to use construction templates. The template provided the same data on facade partitions, walls, roofs, and airtightness. The glazing template was also the same for both buildings. The external window glazing type was project

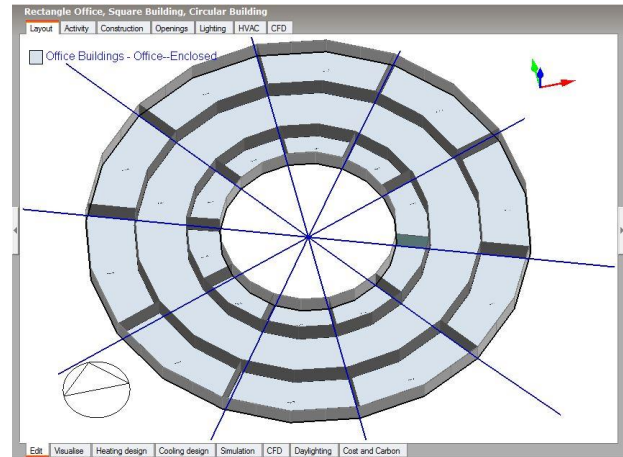


Figure V. 3D plan of the cylinder building modeled by author using DesignBuilder software.

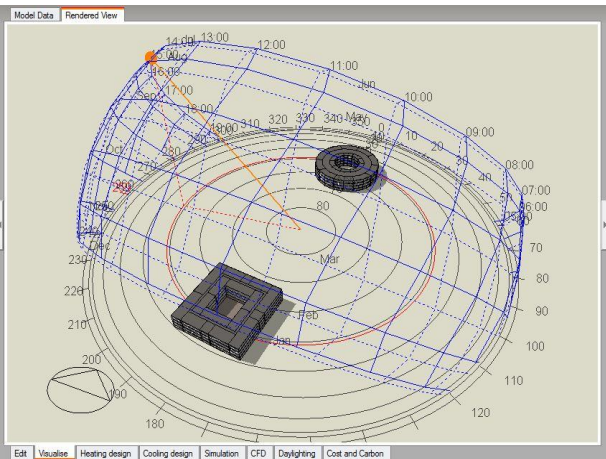


Figure III. 3D view of both cubic and cylinder shapes modeled by author using DesignBuilder software.

externally glazing. And the layout was horizontal strip, with 90% glazed (windows to wall ratio). For the energy simulation process, the energy models were designed to be mechanically ventilated using air conditioning (Fan coil unit

4-

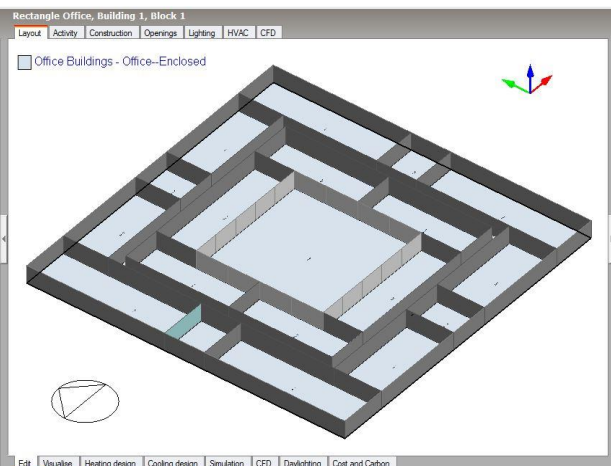


Figure IV. 3D plan of the cubic building modeled by author using DesignBuilder software.

pipe Air-cooled chiller) to calculate the cooling load's

reduction rate. Air changes rate per hour (ac/h) with no fresh air operating pressures (Pa) is left as the DesignBuilder program automatically calculated as 1.7 ac/h @ 50 Pa. DesignBuilder calculates the air rate flow in accordance with ASHRAE standards.

### III. Result and Discussions

The glazing facade is considered to be responsible for a large amount of energy consumption in buildings due to the glass's higher amount of heat exchange as opposed to the other building envelope elements. As mentioned earlier in the energy consumption on buildings section, the overall heat transfer coefficient in  $\text{W/m}^2$  of the building envelope can significantly reduce the energy consumption.

The proposed buildings in this study are fully glazed with vertical solar shading elements, as shown in Figure III. Therefore, internal heat gain through the wall and roof was ignored. The solar heat gain through exterior windows considers as one of the main influence factors on the cooling

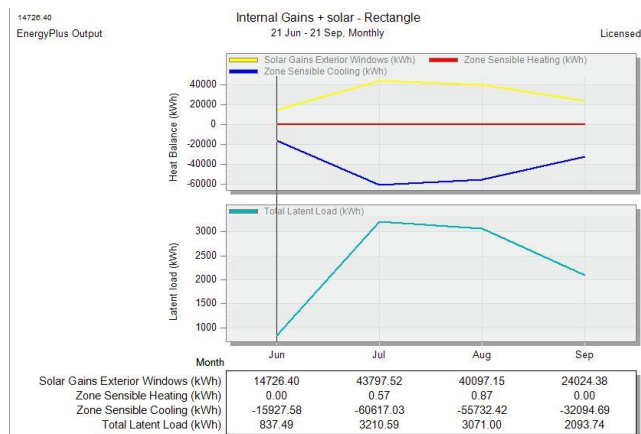


Figure VI. Cubic (rectangular plan) shape building simulation result: showing the impact of solar gain through the exterior windows on cooling load

load of the building. Therefore, the main comparative parameters of the two different forms are indoor operative temperature and solar heat gain through exterior windows. The simulation runs in this step to illustrate the effect of using cubic building form on the amount of solar heat gain

through exterior windows and consequently on indoor operative temperature

. As shown in Figure VI, solar heat gain through the cubic form exterior windows during the peak period in July (43797.52 kWh) has a significant impact on zone sensible cooling load recording (-60617,03 kWh).

Comparing to cylinder form as in figure VII. solar heat gain through the exterior windows of the cylinder form during the peak period at July recorded (38198.14 kWh) lead to a significant impact on zone sensible cooling load recording (-5288,91 kWh).

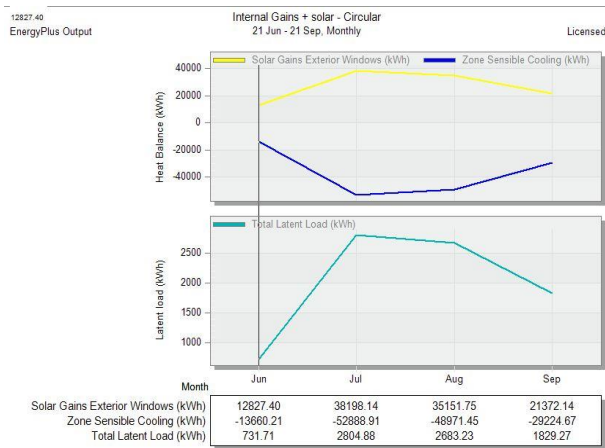


Figure VII. Cylinder (circular plan) form building simulation result: showing the impact of solar gain through the exterior windows on cooling load

Based on these results, the peak cooling load of the cylinder form during summer was recorded (-52888.91 kWh) in July month while its recorded (-60617.03 kWh) in the cubic form in the same month. Therefore, the cylinder form approximately recorded 7728.12 kWh less than the cubic form. According to figures VIII and IX, there is a noticeable effect in lowering indoor operative temperature about 0.2 ~ 0.8 in the cylinder building compared with the cubic building. During the peak cooling load in July, the indoor operative temperature in July in the cubic form recorded 29,27 while it was 29,00 in the cylinder building. And 29,03 at Augustus in the cubic form higher than Cylinder form which recorded 28,73.

To explain the simulation result more clearly, Table I shows the comparative parameters during the cooling period separately. The amount of solar heat gain in the cubic form was higher than the cylinder form during all analyzed months. This high amount of solar heat gain through exterior windows increased the amount of indoor operative temperature as well.

Finally, although the building is located within the same environmental circumstances and has the same construction parameters, there was a noticeable cooling energy consumption differentiation. The results revealed that the circular form is less exposed to the direct sun than the rectangular form. This lead to a decrease in the cooling energy demanded to cool the indoor spaces. According to this fact, minimizing the large direct exposed glass facade area will significantly reduce the cooling load and thus save energy.

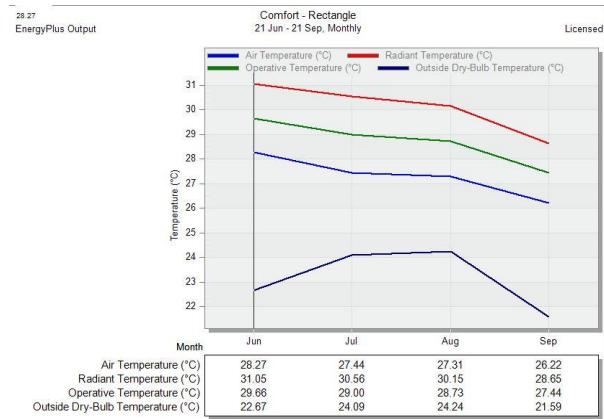


Figure VIII. Cubic (rectangular plan) form building simulation result: showing the impact of solar gain through the exterior windows on indoor operative temperature.

Table I. Comparative analysis results of the two different forms are shown.

Building Form	Cooling period	Solar heat gain exterior windows (kWh)	Operative temperature (C)	Zone sensible cooling load (kWh)
Cylinder (circular plan)	July	38198,14	29.27	-5288,91
Cubic (rectangular plan)		43797.52	29.00	-60617,03
Cylinder (circular plan)	August	3515.75	29.03	-48971.45
Cubic (rectangular plan)		40097.15	28.73	-55732.42
Cylinder (circular plan)	September	21372.14	27.73	-29224.67
Cubic (rectangular plan)		24024.38	27.44	-32094.69

#### IV. Conclusion

In this study, the comparison between two proposed office building cubic (rectangular plan) and cylinder (circular plan) forms in terms of cooling energy consumption was presented. Using DesignBuilder simulation software, building models of the two different forms were modeled to determine the heat gains/ losses and cooling load analysis. Subsequently, the energy monthly cooling loads simulation during the Summer session was conducted to compare the energy cooling consumption on each form. The simulation results showed that cylinder building form received less heat from solar through windows comparing to cubic form. Moreover, the peak cooling load of the cylinder form during summer recorded (-52888.91 kWh) and (-60617.03 kWh) in the cubic form in July, respectively. This indicates that the cylinder form recorded (approximately 7728.12 kWh) less than the cubic form. The presented simulation result shows that when considering building form solutions based on saving energy, obtaining the optimal proportion of the glass area on the building's envelope is a vital solution.

## References

- [1] C. Hachem, A. Athienitis, and P. Fazio, "Parametric investigation of geometric form effects on solar potential of housing units," *Sol. Energy*, vol. 85, no. 9, pp. 1864–1877, Sep. 2011, doi: 10.1016/j.solener.2011.04.027.
- [2] V. N. I. D. M. F. M. F. A. J. Anna Joelsson, "The Impact of the Form Factor on Final Energy Demand in Residential Buildings in Nordic Climates," 2012.
- [3] T. Catalina, J. Virgone, and V. Iordache, "Study on the impact of the building form on the energy consumption," 2011.
- [4] I. E. Kocagil and G. K. Oral, "The effect of building form and settlement texture on energy efficiency for hot dry climate zone in Turkey," in *Energy Procedia*, Nov. 2015, vol. 78, pp. 1835–1840, doi: 10.1016/j.egypro.2015.11.325.
- [5] R. Fallahrafti and M. Mahdavejad, "Optimisation of building Form and orientation for better energy efficient architecture," *Int. J. Energy Sect. Manag.*, vol. 9, no. 4, pp. 593–618, Nov. 2015, doi: 10.1108/IJESM-09-2014-0001.
- [6] A. AlAnzi, D. Seo, and M. Krarti, "Impact of building Form on thermal performance of office buildings in Kuwait," *Energy Convers. Manag.*, vol. 50, no. 3, pp. 822–828, Mar. 2009, doi: 10.1016/j.enconman.2008.09.033.
- [7] B. Bektas Ekici and U. T. Aksoy, "Prediction of building energy needs in early stage of design by using ANFIS," *Expert Syst. Appl.*, vol. 38, no. 5, pp. 5352–5358, May 2011, doi: 10.1016/j.eswa.2010.10.021.
- [8] G. K. Oral and Z. Yilmaz, "Building form for cold climatic zones related to building envelope from heating energy conservation point of view," *Energy Build.*, vol. 35, no. 4, pp. 383–388, May 2003, doi: 10.1016/S0378-7788(02)00111-1.
- [9] V. Bazjanac and L. Berkeley, "Space boundary requirements for modeling of building geometry for energy and other performance simulation," in *Proceedings of the CIB W78 2010: 27th International Conference*, 2010, no. November.
- [10] B. Raof, "THE CORRELATION BETWEEN BUILDING FORM AND BUILDING ENERGY PERFORMANCE.," *Int. J. Adv. Res.*, vol. 5, no. 5, pp. 552–561, May 2017, doi: 10.21474/ijar01/4145.
- [11] S. Pathirana, A. Rodrigo, and R. Halwatura, "Effect of building Form, orientation, window to wall ratios and zones on energy efficiency and thermal comfort of naturally ventilated houses in tropical climate," *Int. J. Energy Environ. Eng.*, vol. 10, no. 3, pp. 107–120, 2019, doi: 10.1007/s40095-018-0295-3.
- [12] P. McKeen and A. S. Fung, "The effect of building aspect ratio on energy efficiency: A case study for multi-unit residential buildings in Canada," *Buildings*, vol. 4, no. 3, pp. 336–354, 2014, doi: 10.3390/buildings4030336.
- [13] M. Premrov, M. Žigart, and V. Žegarac Leskovar, "Influence of the building Form on the energy performance of timber-glass buildings located in warm climatic regions," *Energy*, vol. 149, pp. 496–504, Apr. 2018, doi: 10.1016/j.energy.2018.02.074.
- [14] A. Al-Saggaf, H. Nasir, and M. Taha, "Quantitative approach for evaluating the building design features impact on cooling energy consumption in hot climates," *Energy Build.*, vol. 211, p. 109802, Mar. 2020, doi: 10.1016/j.enbuild.2020.109802.
- [15] M. Mokrzecka, "Influence of building Form and orientation on heating demand: simulations for student dormitories in temperate climate conditions," doi: 10.1051/e3sconf/20184400117.
- [16] A. Zhang, R. Bokel, A. van den Dobbelsteen, Y. Sun, Q. Huang, and Q. Zhang, "The effect of geometry parameters on energy and thermal performance of school buildings in cold climates of China," *Sustain.*, vol. 9, no. 10, p. 1708, Sep. 2017, doi: 10.3390/su9101708.
- [17] A. / Ashrae and / Iesna Addenda, "ANSI/ASHRAE/IESNA Addenda to ANSI/ASHRAE/IESNA Standard 90.1-2007," 2009.
- [18] C. Marino, A. Nucara, and M. Pietrafesa, "Does window-to-wall ratio have a significant effect on the energy consumption of buildings? A parametric analysis in Italian climate conditions," *J. Build. Eng.*, vol. 13, pp. 169–183, Sep. 2017, doi: 10.1016/j.jobe.2017.08.001.
- [19] A. R. AbouElhamd, K. A. Al-Sallal, and A. Hassan, "Review of core/shell quantum dots technology integrated into building's glazing," *Energies*, vol. 12, no. 6, 2019, doi: 10.3390/en12061058.
- [20] K. J. Kontoleon, "Energy saving assessment in buildings with varying façade orientations and types of glazing systems when exposed to sun," *Int. J. Performability Eng.*, vol. 9, no. 1, 2013.
- [21] A. Bhatia, S. A. R. Sangireddy, and V. Garg, "An approach to calculate the equivalent solar heat gain coefficient of glass windows with fixed and dynamic shading in tropical climates," *J. Build. Eng.*, vol. 22, pp. 90–100, Mar. 2019, doi: 10.1016/j.jobe.2018.11.008.
- [22] E. Graiz and W. Al Azhari, "Energy Efficient Glass: A Way to Reduce Energy Consumption in Office Buildings in Amman (October 2018)," *IEEE Access*, vol. 7, 2019, doi: 10.1109/ACCESS.2018.2884991.
- [23] M. Khoukhi, "The combined effect of heat and moisture transfer dependent thermal conductivity of polystyrene insulation material: Impact on building energy performance," *Energy Build.*, vol. 169, pp. 228–235, Jun. 2018, doi: 10.1016/j.enbuild.2018.03.055.
- [24] O. Algburi and F. Behan, "Cooling load reduction in a single-family house, an energy-efficient approach," *Gazi Univ. J. Sci.*, vol. 32, no. 2, pp. 385–400, 2019.
- [25] O. Algburi and F. Beyhan, "Climate-responsive strategies in vernacular architecture of Erbil city," 2019, doi: 10.1080/00207233.2019.1619324.
- [26] F. Fiorito, A. Cannavale, and M. Santamouris, "Development, testing and evaluation of energy savings potentials of photovoltachromic windows in office buildings. A perspective study for Australian climates," *Sol. Energy*, vol. 205, pp. 358–371, Jul. 2020, doi: 10.1016/j.solener.2020.05.080.
- [27] J. Clarke, "Building simulation," in *Energy Simulation in Building Design*, Elsevier, 2001, pp. 64–98.
- [28] W. L. Oberkampf, S. M. Deland, B. M. Rutherford, K. V. Diegert, and K. F. Alvin, "Error and uncertainty in modeling and simulation," *Reliab. Eng. Syst. Saf.*, vol. 75, pp. 333–357, 2002, Accessed: 29 June, 2021. [Online]. Available: [www.elsevier.com/locate/ress](http://www.elsevier.com/locate/ress).
- [29] R. D. Judkoff, "Validation of building energy analysis simulation programs at the solar energy research institute," *Energy Build.*, vol. 10, no. 3, pp. 221–239, Jan. 1988, doi: 10.1016/0378-7788(88)90008-4.

# Comparison of PID and LQR Controller of Autonomous Underwater Vehicle for Depth Control

Osen Fili Nami  
Department of Electrical Engineering  
Universitas Indonesia  
Depok, Indonesia  
osen.fili@ui.ac.id

Abdul Halim  
Department of Electrical Engineering  
Universitas Indonesia  
Depok, Indonesia  
a.halim@ui.ac.id

Dewi H. Budiarti  
Aerospace Engineer  
BPPT  
Jakarta, Indonesia  
dewi.habsari@bppt.go.id

**Abstract**—Autonomous Underwater Vehicle (AUV) is a small unmanned underwater vehicle that is important for Indonesia as an archipelagic country. Apart from military purposes, it is also needed for civilian purposes. For this reason, the development of AUV technology is necessary and has a strategic value. One that should be developed is an AUV dynamic control technology. In this paper, we have designed an AUV depth control model with proportional integral derivative (PID) and linear quadratic regulator (LQR) controllers. A mathematical model of the AUV focused on the depth model has been developed and its stability was analyzed. In the AUV depth model system, an unstable step response was produced, so it is necessary to add a comparison gain to strengthen the stability of the system. Furthermore, the PID and LQR have been designed and were simulated on MATLAB. Eventually, the results of the PID and LQR controllers from the AUV depth control have been analyzed for comparison.

**Keywords**—AUV, Depth Model, Linear Quadratic Regulator, Proportional Integral Derivative, MATLAB

## I. Introduction

Indonesia is an archipelagic country with a total of 17,508 islands including large and small islands, which are stated in Law no. 6 of 1996 concerning Indonesian waters. Based on these data, the islands are separated by waters with a total area of 6,400,000 km<sup>2</sup> [1]. With the vastness of Indonesian waters, there is a great need for equipment that supports surveillance and security defense in all waters in addition to the military submarines already owned. Submarines can only operate in deep waters and their operation is also limited only for the purposes of supervision, defense and security in Indonesian waters. For this reason, mini submarines and unmanned underwater vehicle are needed.

The advantage of unmanned underwater vehicles is that they can be used in shallow waters and inter-island waters that have small gaps. Unmanned submarines are also very efficient and safer for human safety in their operations because they use less manpower. Besides from being able to assist surveillance and security in the waters, it can also be used for civilian purposes such as aquaculture, mapping, mining, and various underwater biota research.

In developing unmanned submarine technology, we must first understand the submarine dynamic control technology. The dynamic model of AUV is based on REMUS [2]. The PID-controller is the most common type of feedback controller throughout the industry [3]. PID and LQR controllers have been introduced in previous studies [4,5].

This paper explains the design of PID and LQR AUV controls and the results were compared.

## II. Modeling of AUV

To design a mathematical model of dynamic motion of a submarine, we must first study the mechanism of motion of the AUV itself. There are two frames of reference to see the position and orientation of the AUV, namely Body Fixed Frame (BFF) and Ear Fixed Frame (EFF) which can be seen in Figure I.

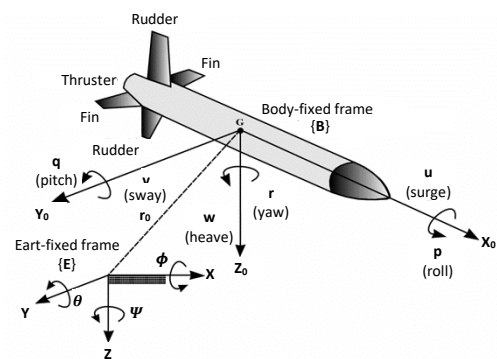


Figure I. AUV reference (body fixed frame and ear fixed frame)

The equations of motion for a AUV are similar to those for a surface ship, however they include all six degrees of freedom [6]. Figure I shows the notations that will be explained in the following Table I.

Tabel I. Notation of AUV axis

AUV Motion	Position	Velocity	Force/Moment
Surge	x	u	X
Sway	y	v	Y
Heave	z	w	Z
Roll	$\phi$	p	K
Pitch	$\theta$	q	M
Yaw	$\psi$	r	N

### A. Model Non-Linear AUV

The motion in the horizontal plane is referred to as surge (longitudinal motion) and sway (sideways motion). Heave is vertical motion. The three DOFs are roll is rotation about the longitudinal axis, pitch is rotation about the transverse axis and yaw is rotation about the vertical axis [7].

Non-linear equation of 6 Degrees of Freedom (6DOF) AUV based on the reference frame of ship's motion is as follows:

Surge:

$$m[\dot{u} - vr + wq - x_G(q^2 + r^2) + y_G(pq - \dot{r}) + z_G(pr + \dot{q})] = \sum X \quad (1)$$

Sway:

$$m[\dot{v} - wp + ur - y_G(r^2 + p^2) + z_G(qr - \dot{p}) + x_G(pq + \dot{r})] = \sum Y \quad (2)$$

Heave:

$$m[\dot{w} - uq + vp - z_G(p^2 + q^2) + x_G(rp - \dot{q}) + y_G(rq + \dot{p})] = \sum Z \quad (3)$$

Roll:

$$I_x \dot{p} + (I_z - I_y)qr - (\dot{r} + pq)I_{xz} + (r^2 - q^2)I_{yz} + (pr - \dot{q})I_{xy} + m[y_G(\dot{w} - uq + vp) - z_G(\dot{v} - wp + ur)] = \sum K \quad (4)$$

Pitch:

$$I_y \dot{q} + (I_x - I_z)rp - (\dot{p} + qr)I_{xy} + (p^2 - r^2)I_{xz} + (pq - \dot{r})I_{yz} + m[z_G(\dot{u} - vr + wq) - x_G(\dot{w} - uq + vp)] = \sum M \quad (5)$$

Yaw:

$$I_z \dot{r} + (I_y - I_x)pq - (\dot{q} + rp)I_{yz} + (q^2 - p^2)I_{xy} + (rq - \dot{p})I_{xz} + m[x_G(\dot{v} - wp + ur) - y_G(\dot{u} - vr + wq)] = \sum N \quad (6)$$

Total force and moment when inserted into the equation 6DOF AUV then the equation will be as below:

Surge:

$$m[\dot{u} - vr + wq - x_G(q^2 + r^2) + y_G(pq - \dot{r}) + z_G(pr + \dot{q})] = X_{res} + X_{|u|u}u|u| + X_{\dot{u}}\dot{u} + X_{wq}wq + X_{qq}qq + X_{vr}vr + X_{rr}rr + X_{prop} \quad (7)$$

Sway:

$$m[\dot{v} - wp + ur - y_G(r^2 + p^2) + z_G(qr - \dot{p}) + x_G(pq + \dot{r})] = Y_{res} + Y_{|v|v}v|v| + Y_{r|r}r|r| + Y_{\dot{v}}\dot{v} + Y_{\dot{r}}\dot{r} + Y_{ur}ur + Y_{wp}wp + Y_{pq}pq + Y_{uv}uv + Y_{uu\delta_r}u^2\delta_r \quad (8)$$

Heave:

$$m[\dot{w} - uq + vp - z_G(p^2 + q^2) + x_G(rp - \dot{q}) + y_G(rq + \dot{p})] = Z_{res} + Z_{|w|w}w|w| + Z_{q|q}q|q| + Z_{\dot{w}}\dot{w} + Z_{\dot{q}}\dot{q} + Z_{uq}uq + Z_{vp}vp + Z_{rp}rp + Z_{uw}uw + Z_{uu\delta_s}u^2\delta_s \quad (9)$$

Roll:

$$I_x \dot{p} + (I_z - I_y)qr - (\dot{r} + pq)I_{xz} + (r^2 - q^2)I_{yz} + (pr - \dot{q})I_{xy} + m[y_G(\dot{w} - uq + vp) - z_G(\dot{v} - wp + ur)] = K_{res} + K_{p|p}p|p| + K_{\dot{p}}\dot{p} + K_{prop} \quad (10)$$

Pitch:

$$I_y \dot{q} + (I_x - I_z)rp - (\dot{p} + qr)I_{xy} + (p^2 - r^2)I_{xz} + (pq - \dot{r})I_{yz} + m[z_G(\dot{u} - vr + wq) - x_G(\dot{w} - uq + vp)] = M_{res} + M_{|w|w}w|w| + M_{q|q}q|q| + M_{\dot{w}}\dot{w} + M_{\dot{q}}\dot{q} + M_{uq}uq + M_{vp}vp + M_{rp}rp + M_{uw}uw + M_{uu\delta_s}u^2\delta_s \quad (11)$$

Yaw:

$$I_z \dot{r} + (I_y - I_x)pq - (\dot{q} + rp)I_{yz} + (q^2 - p^2)I_{xy} + (rq - \dot{p})I_{xz} + m[x_G(\dot{v} - wp + ur) - y_G(\dot{u} - vr + wq)] = N_{res} + N_{|v|v}v|v| + N_{r|r}r|r| + N_{\dot{v}}\dot{v} + N_{\dot{r}}\dot{r} + N_{ur}ur + N_{wp}wp + N_{pq}pq + N_{uv}uv + N_{uu\delta_r}u^2\delta_r \quad (12)$$

By separating the acceleration notation and assumed diagonal inertia tensor ( $I_{xy}, I_{xz}, I_{yz}$ ) is zero then equations 7 to equation 12 will be simplified back to:

Surge:

$$(\mu - X_{\dot{u}})\dot{u} + mZ_G\dot{q} - mY_G\dot{r} = X_{res} + X_{|u|u}u|u| + (X_{wq} - m)wq + (X_{qq} + mx_G)qq + (X_{vr} + m)vr + (X_{rr} + mx_G)rr - mY_Gpq - mZ_Gpr + X_{prop}$$

where:

$$\sum X_{ext} = X_{res} + X_{|u|u}u|u| + (X_{wq} - m)wq + (X_{qq} + mx_G)qq + (X_{vr} + m)vr + (X_{rr} + mx_G)rr - mY_Gpq - mZ_Gpr + X_{prop}$$

So that the equation becomes:

$$(\mu - X_{\dot{u}})\dot{u} + mZ_G\dot{q} - mY_G\dot{r} = \sum X_{ext} \quad (13)$$

Sway:

$$(m - Y_{\dot{v}})\dot{v} - mZ_G\dot{p} + (mx_G - Y_{\dot{r}})\dot{r} = Y_{res} + Y_{|v|v}v|v| + Y_{r|r}r|r| + mY_Gr^2 + (Y_{ur} - m)ur + (Y_{wp} + m)wp + (Y_{pq} - mx_G)pq + Y_{uv}uv + mY_Gp^2 + mZ_Gqr + Y_{uu\delta_r}u^2\delta_r$$

where:

$$\sum Y_{ext} = Y_{res} + Y_{|v|v}v|v| + Y_{r|r}r|r| + mY_Gr^2 + (Y_{ur} - m)ur + (Y_{wp} + m)wp + (Y_{pq} - mx_G)pq + Y_{uv}uv + mY_Gp^2 + mZ_Gqr + Y_{uu\delta_r}u^2\delta_r$$

So that the equation becomes:

$$(m - Y_{\dot{v}})\dot{v} - mZ_G\dot{p} + (mx_G - Y_{\dot{r}})\dot{r} = \sum Y_{ext} \quad (14)$$

Heave:

$$(m - Z_{\dot{w}})\dot{w} + mY_G\dot{p} - (mx_G + Z_{\dot{q}})\dot{q} = Z_{res} + Z_{|w|w}w|w| + Z_{q|q}q|q| + (Z_{uq} + m)uq + (Z_{vp} - m)vp + (Z_{rp} - mx_G)rp + Z_{uw}uw + mZ_G(p^2 + q^2) - mY_Grq + Z_{uu\delta_s}u^2\delta_s$$

where:

$$\sum Z_{ext} = Z_{res} + Z_{|w|w}w|w| + Z_{q|q}q|q| + (Z_{uq} + m)uq + (Z_{vp} - m)vp + (Z_{rp} - mx_G)rp + Z_{uw}uw + mZ_G(p^2 + q^2) - mY_Grq + Z_{uu\delta_s}u^2\delta_s$$

So that the equation becomes:

$$(m - Z_{\dot{w}})\dot{w} + mY_G\dot{p} - (mx_G + Z_{\dot{q}})\dot{q} = \sum Z_{ext} \quad (15)$$

Roll:

$$-mZ_G\dot{v} + mY_G\dot{w} + (I_x - K_{\dot{p}})\dot{p} = K_{res} + K_{p|p}p|p| - (I_z - I_y)qr + m(uq - vp) - mZ_G(wp - ur) + K_{prop}$$

where:

$$\sum K_{ext} = K_{res} + K_{p|p}p|p| - (I_z - I_y)qr + m(uq - vp) - mZ_G(wp - ur) + K_{prop}$$

So that the equation becomes:

$$-mZ_G\dot{v} + mY_G\dot{w} + (I_x - K_{\dot{p}})\dot{p} = \sum K_{ext} \quad (16)$$

Pitch:

$$mZ_G\dot{u} - (mx_G + M_{\dot{w}})\dot{w} + (I_y - M_{\dot{q}})\dot{q} = M_{res} + M_{|w|w}w|w| + M_{q|q}q|q| + (M_{uq} - mx_G)uq + (M_{vp} + mx_G)vp + (M_{rp} - (I_x - I_z))rp + mZ_G(vp - wp) + M_{uw}uw + M_{uu\delta_s}u^2\delta_s$$

where:

$$\sum M_{ext} = M_{res} + M_{|w|w}w|w| + M_{q|q}q|q| + (M_{uq} - mx_G)uq + (M_{vp} + mx_G)vp + (M_{rp} - (I_x - I_z))rp + mZ_G(vp - wp) + M_{uw}uw + M_{uu\delta_s}u^2\delta_s$$

So that the equation becomes:

$$mz_G\dot{u} - (mx_G + M_w)\dot{w} + (I_y - M_q)\dot{q} = \Sigma M_{ext} \quad (17)$$

Yaw :

$$\begin{aligned} -my_G\dot{u} + (mx_G - N_v)\dot{v} + (I_z - N_r)\dot{r} &= N_{res} + \\ N_{|v|v}v|v| + N_{r|r}r|r| + (N_{ur} - mx_G)ur &+ (N_{wp} + \\ mx_G)wp + (N_{pq} - (I_y - I_x))pq - my_G(vr - wq) &+ \\ N_{uv}uv + N_{uu\delta_r}u^2\delta_r \end{aligned}$$

where:

$$\begin{aligned} \Sigma N_{ext} &= N_{res} + N_{|v|v}v|v| + N_{r|r}r|r| + (N_{ur} - \\ mx_G)ur + (N_{wp} + mx_G)wp + (N_{pq} - (I_y - I_x))pq - \\ my_G(vr - wq) + N_{uv}uv + N_{uu\delta_r}u^2\delta_r \end{aligned}$$

So that the equation becomes:

$$-my_G\dot{u} + (mx_G - N_v)\dot{v} + (I_z - N_r)\dot{r} = \Sigma N_{ext} \quad (18)$$

## B. Linearisasi Model Non-Linear Depth Model

It is assumed that the dept-plane motion is in ideal condition so that we only need the surge velocity, heave velocity and pitch rate as well as the forward position  $x$ ,  $z$  depth and pitch  $\theta$  angle. It will be assumed that velocity other than that will be created equal to zero ( $v, p, r$  dan  $y_g$ ). Heave ( $w$ ) and pitch ( $q$ ) linearized to zero so that equations 1, 2 and 5 become :

$$\Sigma X = m(\dot{u} + z_G\dot{q}) \quad (19)$$

$$\Sigma Z = m(\dot{w} - uq - x_G\dot{q}) \quad (20)$$

$$\Sigma M = I_y\dot{q} + m[z_G\dot{u} - x_G(\dot{w} - uq)] \quad (21)$$

The total depth-plane force and moments shown in equations 22,23 and 24 below:

$$\Sigma X = X_u\dot{u} + X_uu + X_qq + X_\theta\theta \quad (22)$$

$$\Sigma Z = Z_w\dot{w} + Z_q\dot{q} + Z_w w + Z_q q + Z_{\delta_s}\delta_s \quad (23)$$

$$\Sigma M = M_w\dot{w} + M_q\dot{q} + M_w w + M_q q + M_\theta\theta + M_{\delta_s}\delta_s \quad (24)$$

From equations 19-21 and from equations 22-24 the linear equation of the depth model AUV motion becomes:

$$(m - X_u)\dot{u} + mz_G\dot{q} - X_uu - X_qq - X_\theta\theta = 0 \quad (25)$$

$$\begin{aligned} (m - Z_w)\dot{w} - (mx_G + Z_q)\dot{q} - Z_w w \\ - (mU + Z_q)q = Z_{\delta_s}\delta_s \end{aligned} \quad (26)$$

$$\begin{aligned} mz_G\dot{u} - (mx_G + M_w)\dot{w} + (I_y - M_q)\dot{q} - \\ M_w w + (mx_G U - M_q)q - M_\theta\theta = M_{\delta_s}\delta_s \end{aligned} \quad (27)$$

Then it is assumed that  $z_G$  is very small value compared to others so that heave and pitch can be decouple. from the surge equation resulting in motion equations as follows:

$$\begin{aligned} (m - Z_w)\dot{w} - (mx_G + Z_q)\dot{q} - \\ Z_w w - (mU + Z_q)q = Z_{\delta_s}\delta_s \end{aligned} \quad (28)$$

$$\begin{aligned} - (mx_G + M_w)\dot{w} + (I_y - M_q)\dot{q} - \\ M_w w + (mx_G U - M_q)q - M_\theta\theta = M_{\delta_s}\delta_s \end{aligned} \quad (29)$$

Then linearize kinematic equations between eart and body-fixed velocities:

$$\dot{z} = -U\theta + w \quad (30)$$

$$\dot{\theta} = q \quad (31)$$

From equations 28 and 29 and equations 30 and 31 obtain simplified equations in the matrix below:

$$\begin{bmatrix} m - Z_w & -(mx_G + Z_q) & 0 & 0 \\ -(mx_G + M_w) & I_y - M_q & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \dot{w} \\ \dot{q} \\ \dot{z} \\ \dot{\theta} \end{bmatrix} - \begin{bmatrix} Z_w & mU + Z_q & 0 & 0 \\ M_w & -mx_G U + M_q & 0 & M_\theta \\ 1 & 0 & 0 & -U \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} w \\ q \\ z \\ \theta \end{bmatrix} = \begin{bmatrix} Z_{\delta_s} \\ M_{\delta_s} \\ 0 \\ 0 \end{bmatrix} [\delta_s] \quad (32)$$

From equation 32 it is assumed  $x_G = 0$ , the heave velocity is assumed always small so that both  $w$  and  $\dot{w}$  can be ignored so the equation 32 becomes:

$$\begin{bmatrix} I_y - M_q & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \dot{q} \\ \dot{z} \\ \dot{\theta} \end{bmatrix} - \begin{bmatrix} M_q & 0 & M_\theta \\ 0 & 0 & -U \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} q \\ z \\ \theta \end{bmatrix} = \begin{bmatrix} M_{\delta_s} \\ 0 \\ 0 \end{bmatrix} [\delta_s] \quad (33)$$

Will be regretted:

$$\text{matrix} \begin{bmatrix} I_y - M_q & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = T1$$

$$\text{matrix} \begin{bmatrix} M_q & 0 & M_\theta \\ 0 & 0 & -U \\ 1 & 0 & 0 \end{bmatrix} = T2 \text{ dan}$$

$$\text{matrix} \begin{bmatrix} M_{\delta_s} \\ 0 \\ 0 \end{bmatrix} = T3$$

state becomes  $x = [q \quad z \quad \theta]^T$

input  $u = [\delta_s]^T$

output,  $y = z$

$$T1\dot{x} - T2x = T3u \quad (34)$$

$$\dot{x} = T1^{-1}T2x + T1^{-1}T3u \quad (35)$$

So if transferred to the form of state space

Matrix  $A=T1^{-1}T2$  and matrix  $B=T1^{-1}T3$

## C. Parameter-parameter AUV

All AUV parameter values to be used in this study are using REMUS unmanned submarine model parameters. With the forward speed of the ship is  $u=1.54$  m/s and the mass of the submarine is 30 kg. The other parameters are in Table II below.

Table II. REMUS AUV Parameters

Parameter	Value	Units	Description
$I_x$	+1.77e-001	kg.m <sup>2</sup>	M.I w.r.t Origin at CB
$I_y$	+3.45e+000	kg.m <sup>2</sup>	M.I w.r.t Origin at CB
$I_z$	+3.45e+000	kg.m <sup>2</sup>	M.I w.r.t Origin at CB
$Z_q$	-9.67e+000	kg.m/s	Combined Term
$Z_{\dot{q}}$	-1.93e+000	kg.m	Added Mass
$Z_w$	-6.66e+001	kg/s	Combined Term
$Z_{\dot{w}}$	-3.55e+001	kg	Added Mass
$Z_{\delta_s}$	-5.06e+001	kg.m/s <sup>2</sup>	Fin Lift
$M_q$	-6.87e+000	kg.m <sup>2</sup> /s	Combined Term
$M_{\dot{q}}$	-4.88e+000	kg.m <sup>2</sup>	Added Mass
$M_w$	+3.07e+001	kg.m/s	Combined Term
$M_{\dot{w}}$	-1.93e+000	kg.m	Added Mass
$M_{\delta_s}$	-3.46e+001	kg.m <sup>2</sup> /s <sup>2</sup>	Added Mass
$M_{\theta}$	-5.77e+000	kg.m <sup>2</sup> /s <sup>2</sup>	Hydrostatic

### D. Depth Model AUV

Model State Space

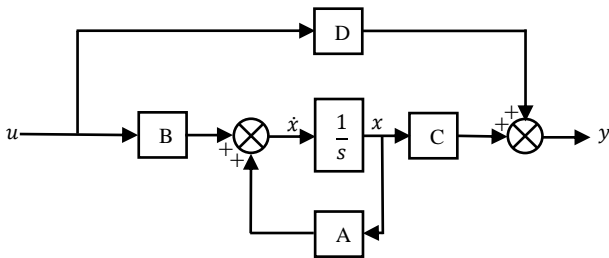


Figure II. Blok Diagram State Space

State Space Equation from Figure II:

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t) \\ y(t) &= Cx(t) + Du(t) \end{aligned} \quad (36)$$

State space Depth Model :

After entering the parameters in Table II into equation 33 obtained:

$$\dot{x} = Ax + Bu$$

$$\text{State : } \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} q \\ z \\ \theta \end{bmatrix}$$

$$\text{Input : } u = [\delta_s]^T$$

$$\text{Output : } y = [z]$$

Matrix

$$T1 = \begin{bmatrix} 8.33 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$T1^{-1} = \begin{bmatrix} 0.12 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$T2 = \begin{bmatrix} -6.87 & 0 & -5.77 \\ 0 & 0 & -1.54 \\ 1 & 0 & 0 \end{bmatrix}$$

$$T3 = \begin{bmatrix} -34.6 \\ 0 \\ 0 \end{bmatrix}$$

So the matrix

$$A = T1^{-1}T2$$

$$A = \begin{bmatrix} -0.8247 & 0 & -0.6927 \\ 0 & 0 & -1.5400 \\ 1 & 0 & 0 \end{bmatrix}$$

and

$$B = T1^{-1}T3$$

$$B = \begin{bmatrix} -4.1537 \\ 0 \\ 0 \end{bmatrix}$$

While the matrix

$$C = [0 \quad 1 \quad 0]$$

and

$$D = [0]$$

So that the state space matrix becomes:

$$\dot{x}(t) = Ax(t) + Bu(t)$$

$$\begin{bmatrix} \dot{q} \\ \dot{z} \\ \dot{\theta} \end{bmatrix} = \begin{bmatrix} -0.8247 & 0 & -0.6927 \\ 0 & 0 & -1.5400 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} q \\ z \\ \theta \end{bmatrix} + \begin{bmatrix} -4.1537 \\ 0 \\ 0 \end{bmatrix} [\delta_s]$$

$$y(t) = [0 \quad 1 \quad 0] \begin{bmatrix} q \\ z \\ \theta \end{bmatrix} + [0] [\delta_s]$$

$$y(t) = z \quad (37)$$

Transfer function Depth Model becomes:

$$G(s) = \frac{6.397}{s^3 + 0.8247s^2 + 0.6927s} \quad (38)$$

### E. AUV System Depth Model Stability Analysis

From the Transfer function Depth open loop model can be determined the root position of the system. The root locus can be observed in Figure III.

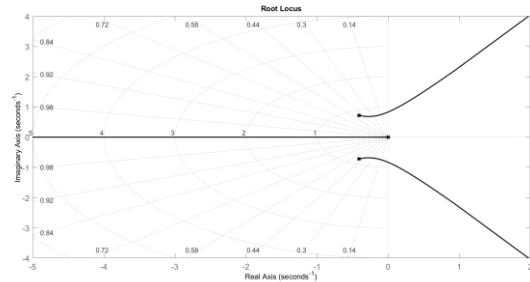


Figure III. Root Locus Depth model

Zero and pole of depth model are as follows:

Zero : none

Pole 1 : 0.0000 + 0.0000i

Pole 2 : -0.4123 + 0.7230i

Pole 3 : -0.4123 - 0.7230i

When viewed from the root position two poles are on the left axis imaginer and one pole is right in the zero position should the system can be said to be stable. However, further analysis of system stability is needed by looking at bode diagram and step response system.



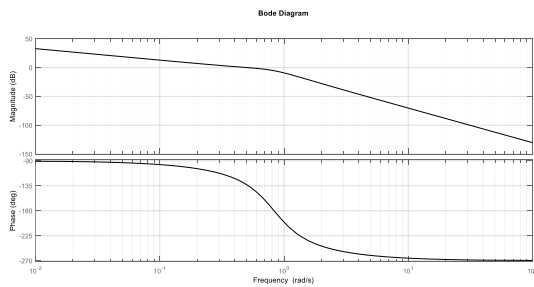


Figure IV. Diagram Bode Depth Model

From bode Gain Margin System diagram of 0.0893, Phase Margin -61.9435 of Negative margin phase can also be concluded that the system is unstable.

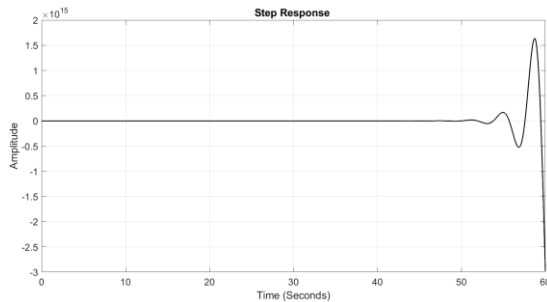


Figure V. Close Loop Step Respon Depth Model

From the V images can be seen step response of the unstable depth model. Therefore, it is necessary to add Gain Compensator to compensate for system instability. From root locus we can specify the Gain to be used. To determine the Compensator Gain can be seen for example in the literature [8]. In this paper used Gain of 0.048 Then the Gain is added to the system.

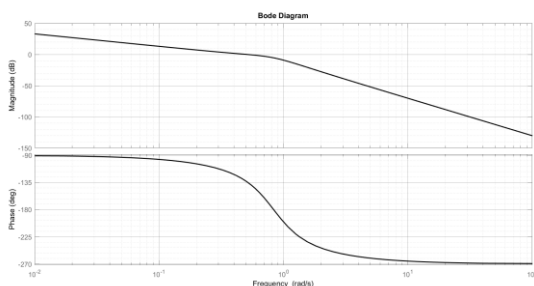


Figure VI. Bode Depth Model diagram added Gain Compensator

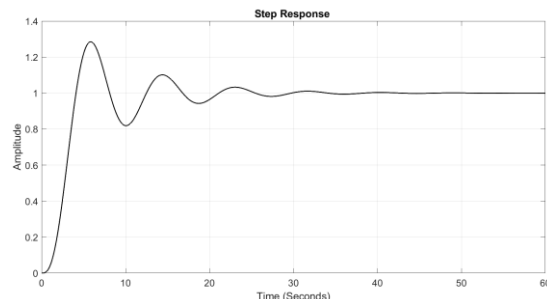


Figure VII. Close Loop Step Response Depth Model with Gain Compensator.

From bode diagram contained in Figure VI the gain margin value is 1.8608 and the phase margin is 45.9785. From this can be said stable system. The Close Loop Step

response Depth Model with gain change obtained graph on figure VII with Rise time of 2.4036, Settling time of 24.3448, Overshoot of 28.4421 and steady state error of 0.027. From the step response in figure VII although the system is able to stabilize but still needed controller to be able to better stabilize the system with better conditions.

### III. Control Methods and Simulation

The depth model will be designed with PID control and LQR controller to help stabilize the system.

#### A. PID Control Model

The PID Control equation is :

$$G_c(s) = K_p \left( 1 + \frac{1}{T_i s} + T_d s \right) \quad (39)$$

To guide the gain of  $K_p$ ,  $T_i$  and  $T_d$  on depth control the model can use the Ziegler-Nicholes method. To help gain tuningnya can be seen MATLAB Code Gain Tunning Ziegler-Nicholes as introduced by Dingyu et al. (2009) [9].

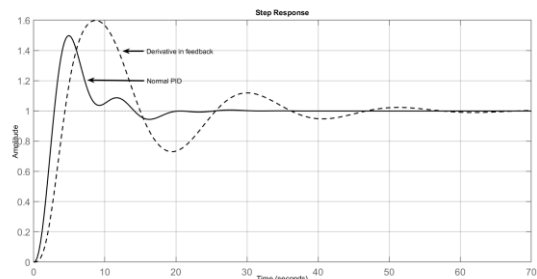


Figure VIII. Depth Control Normal PID response step and with Derivative in feedback

The controller designed are :

$$G_{PID} = 1.1165 \left( 1 + \frac{1}{3.7747s} + 0.9059s \right)$$

With value of  $K_p=1.1165$ ,  $T_i=3.7747$  dan  $T_d=0.9059$

From the performance of step response depth model controller using Normal PID is rise time 1.8386 seconds, Settling time 18.5743 seconds, overshoot 49.798, Peak 1.4980, peak time 5.0172 seconds and steady state errornya 0.0017.

Step response performance from PID controller with derivative in the feedback path is rise time 2.55 seconds, settling time 53.0766 seconds, overshootnya 59.8747, Peak 1.5987, peak time 8.6876 seconds and steady state error 7.5162e-04.

From the results of the response step of the two PID controllers at Figure VIII, both are able to control the depth of the system model well. Although steady state error for PID controller with derivative in the feedback path is smaller even hamper near zero but in terms of performance is better performance of ordinary PID Controller compared to PID with derivative in the feedback path.

## B. LQR Control Model

LQR is one of the optimal control methods based on state space and is essentially intended to determine the control signal in such a way as to minimize the performance index  $J$ .

$$J = \int_0^{\infty} (x^* Q x + u^* R u) dt \quad (40)$$

To minimize the  $J$  performance index of the equation 40 of course we must first determine the values  $Q$  and definite.  $Q$  matrix is the weighting matrix state and  $R$  is the weighting matrix of control and both are definite and Hermitian positive matrices. Before being inserted into equation 40, the values  $Q$  and  $R$  are used first to produce a definite positive matrix of  $P$  by inserting it into equation 41 namely Algebraic Riccati Equation (ARE). After  $P$  is produced it will be used to look for the optimal  $K$  booster matrix contained in equation 42. And once obtained matrix  $K$  will be obtained control signal whose equation can be seen in equation 43.

$$A^* P + P A - P B R^{-1} B^* P + Q = 0 \quad (41)$$

$$K = R^{-1} B^* P \quad (42)$$

$$u = -Kx \quad (43)$$

To facilitate in the calculation of LQR control can use matlab code on reference [9].

To determine the values  $R$  and  $Q$  can be seen [10]. In depth model LQR is given matrix  $Q$  and  $R$  are as follows:

$$Q = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix} \text{ and } R = [1]$$

Then inserted into the equation 46 obtained matrix values:

$$P = \begin{bmatrix} 0.3089 & -0.2407 & 0.5779 \\ -0.2407 & 1.6671 & -1.4817 \\ 0.5779 & -1.4817 & 3.4001 \end{bmatrix} \text{ and}$$

$$K = [-1.2831 \quad 1.0000 \quad -2.4006]$$

The eigen value can also be calculated by MATLAB code  $eig(A-B*K)$  obtained values as follows:

$$-3.7818 + 0.0000i$$

$$-1.1863 + 0.5331i$$

$$-1.1863 - 0.5331i$$

It can be seen that the eigen value is all negative so that its position is to the left of the imaginary axis and it is concluded that the system is stable.

Then the response step of LQR control for depth control can be seen in Figure IX.

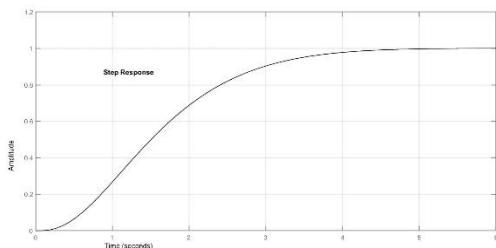


Figure IX. Step response LQR Depth Model

In Figure IX obtained excellent response step with rise time of 2.362 seconds, Settling time of 4.0418 seconds, overshoot 0.0836, Peak 1.0008, peak time of 6.2888 seconds and steady state error of  $5.8645e-04$ . It can be concluded that

the LQR controller design is very good at stabilizing the depth of the AUV model.

## C. Comparison of PID and LQR controller results

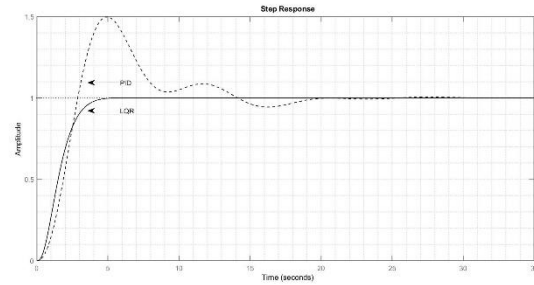


Figure X. Step response comparison graph of PID control and LQR control on Depth Model.

From Figure X, it can be clearly seen that the control result with optimal control of LQR is better performance compared to PID control. This can be seen also with step response data, namely rise time, settling time, overshoot and steady state error from both controllers can be concluded that a better depth control model is to use optimal LQR control.

## IV. Conclusion

From the simulation of PID control design and LQR Depth Control Model both are able to control the AUV Depth Model well. However, when compared between the two, LQR controls gives better performance compared to PID control, especially observing from the settling time and overshoot. LQR control has faster settling time of 4.0418 seconds and very small overshoot of 0.0836 compared to PID controller which has settling time of 18.5743 seconds and overshoot of 49.798.

## References

- [1] kkp.go.id.(2019, 3 November). Menko Maritim Luncurkan Data Rujukan Wilayah Kelautan Indonesia. <https://kkp.go.id/brsdm/poltekkarawang/artikel/14863-menko-maritim-luncurkan-data-rujukan-wilayah-kelautan-indonesia>.
- [2] Timothy Prester, "Verification of Six-Degree of Freedom Simulation Model for REMUS Autonomous Underwater Vehicle", Massachusetts Institute Of Technology and the Woods Hole Oceanographic Institution, M.Sc Thesis, 2001.
- [3] Erik Lind, Magnus Meijer, "Simulation and Control of Submarines", Lund University, M.Sc Thesis, 2014.
- [4] Katsuhiko Ogata, Modern Control Engineering Fifth Edition , Prentice Hall, New Jersey, 2010.
- [5] Lectures, "Linear Quadratic Regulator (LQR) State Feedback Design", F.L.Lewis 1998, Update: Monday, September 16, 2019.
- [6] Martin Renilson, Submarine Hydrodynamics, SpringerBriefs in Applied Sciences and Technology, 2015.
- [7] Thor I. Fossen, " A Nonlinear Unified State-Space Model For Ship Maneuvering And Control in Seaway", Journal of Bifurcation and Chaos, vol.15, no. 9, pp. 2717-2746, 2005.
- [8] Norman S. Nise, Control System Engineering Sixth Edition, John Wiley&Sons,Inc, United States of America, 2011.
- [9] Dingyu Xue, TangQuan Chen, Derek P.Atherton, Linear Feedback Control Analysis and Design with MATLAB, Society for Industrial and Applied Mathematics, Philadelphia, 2007.
- [10] Yul Y.Nazaruddin, Franky, I.G.N.A.Indra Mandala, "Optimasi Pengontrol LQR menggunakan Algoritma Stochastic Fractal Search", Seminar Nasional Instrumentasi, Kontrol dan Otomasi (SNIKO), Bandung, Indonesia, 10-11 Desember 2018
- [11] www.mathworks.com

# Convolutional Neural Network Approach to Distinguish and Characterize Tumor Samples Using Gene Expression Data

Büşra Nur Darendeli  
Department of Bioengineering  
Yildiz Technical University  
Istanbul, Turkey  
[bndarendeli@gmail.com](mailto:bndarendeli@gmail.com)

Alper Yilmaz  
Department of Bioengineering  
Yildiz Technical University  
Istanbul, Turkey  
[alyilmaz@yildiz.edu.tr](mailto:alyilmaz@yildiz.edu.tr)

**Abstract**— Cancer is threatening millions of people each year and its early diagnosis is still a challenging task. Early diagnosis is one of major ways to tackle this disease and lower the mortality rate. Advancements in deep learning approaches and availability of biological data offer applications that can facilitate the diagnosis and characterization of cancer. Here, we aimed to provide new perspective of cancer diagnosis using deep learning approach on gene expression data. We turn the information of gene expression data of cancer and normal tissues into input for Convolutional Neural Network (CNN) with the method we call RGB mapping. It is aimed to be learned with CNN by preserving the gene expression data values of the tissues. In this way, it is aimed to protect the effect of each gene in cancer diagnosis. In addition, we aimed to characterize the disease by identifying genes that are effective in cancer prediction. In this study, The Cancer Genome Atlas (TCGA) dataset with RNA-Seq data of approximately 30 different types of cancer patients and GTEx RNA-seq data of normal tissues were used. The input data for the training was transformed to RGB format and the training was carried out with a CNN. The trained algorithm is able to predict cancer with 97,7% accuracy, based on gene expression data. Moreover, we applied one-pixel attack on the trained model to determine effective genes for prediction of the disease. As a result of this attack, 13 genes that are effective on the decision mechanism of the algorithm were determined. In conclusion, a new data preprocessing method is proposed before the training in this study. With the one-pixel attack method, which can be applied to the model trained using this method, it has been possible to identify genes that may be biomarkers in cancer. The effects of candidate genes on cancer can be determined by experimental studies.

**Keywords**— cancer, CNN, TCGA, GTEx, RNA-seq, RGB Mapping

## I. Introduction

The deep learning approach has emerged by designing computer models that can perform the learning process as a result of interconnected layers based on the human brain, such as neurons. As a result of the development of data science and especially the rapid increase in biological data in the last decade, the designed neural networks have begun to play important roles in the interpretation of biological data for the diagnosis and treatment of diseases [1]. Cancer, which is one of the biggest health problems in the world, is one of the diseases in which deep learning approaches are widely applied.

Since cancer is a disease with high genomic heterogeneity and phenotypic plasticity, its diagnosis and treatment involve various difficulties [2]. Thanks to the developing technology, many medical data of cancer patients are available. As a result of the processing of these medical data with deep learning approaches, the stages of diagnosis and treatment have improved.

The presence of CT and histopathology data in the diagnosis of the disease enabled improvements in the diagnosis stage as a result of image-based processing using the deep learning approach of these data. Deep learning algorithms have been used in the diagnosis of many types of cancer, including breast cancer [3,4], prostate cancer [5,6], lung cancer [7,8], colon cancer[9], head and neck cancer [10], and skin cancer[11]. These studies, which are based on image processing, are widely used as they provide advantages to clinicians at the early diagnosis stage.

In addition to the image-based approach, biological data have also been used in cancer diagnosis [12] and even treatment [13] approaches. Gene expression signature-based approach is generally used in eliminating the disadvantages caused by heterogeneity at the diagnosis stage. Gene expression data and deep learning approach are used for many methods such as estimation of survival times of individuals with cancer [14], determination of biomarker genes [15], classification [16,17]. All of these studies show that by using gene expression data and deep learning approaches together, important information will be gained about the mechanism of cancer.

In this study, The Cancer Genome Atlas (TCGA) dataset with RNA-seq data of approximately 30 different types of cancer patients and a dataset obtained by curation of GTEx data including RNA-seq analysis of normal tissues was used. Importance was given to homogeneity of tumor and normal tissue distribution of the prepared dataset. The gene expression values of the tissues, which are the input data for training, were converted into 24-bit binary format. Then it was converted to 8 bits of Red, Green and Blue channels. The training was carried out with the CNN algorithm. The trained algorithm is able to diagnose cancer and normal patients 97.7% with accuracy, based on gene expression data. Afterwards, one-pixel attack method was applied to the input data created using RGB mapping. In this way, the vulnerability of deep learning models has been used to identify genes that may be effective in cancer. As a result of this process, 13 biomarker genes were obtained. When these genes were investigated in the literature, their relationship with cancer was determined.

## II. Materials and Methods

### A. Dataset Preparation

Data downloaded from UCSC Xena platform [18] which includes three different RNA-Seq data sources; TARGET, TCGA and GTEx. Dataset label distribution is shown in Table I.

Table I. Distribution of labels in whole dataset.

Datasets	Normal	Tumor
TCGA	727	9750
GTEX	7429	0

Data labels (Normal, Tumor) were extracted from phenotype information of selected samples. Gene IDs were converted from Entrez ID to ENSEMBL IDs using BioMart online tool. The differentially expressed gene list [19] was used to select 1024 genes which show highest up-regulation or down-regulation count throughout the whole dataset. Expression data for selected genes showing the most expressivity value were used as input for training.

### B. Proposed Data Preprocessing Method: RGB Mapping

Gene expression values were converted into (R,G,B) format before the trained step. RGB values are obtained by converting gene expression value into 24 bit long binary and then using first 8 bits for R (red), second 8 bits for G (green) and third 8 bits for B (blue) (Figure I). For each sample, 32x32x3 3D Numpy array was prepared .

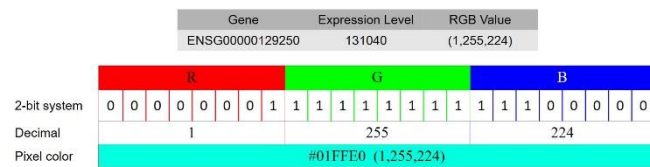


Figure I. Conversion of gene expression value to RGB format

Figure II summarizes the concept with an sample conversion. After the conversion, 14,308 x 1,024 training data becomes [14,308 x 32 x 32 x 3] Numpy array. Accordingly, test data becomes [3,578 x 32 x 32 x 3] Numpy array suitable for batch processing by Tensorflow. Final layer represents R,G,B values, the square 32x32 shape corresponds to 1,024 genes.

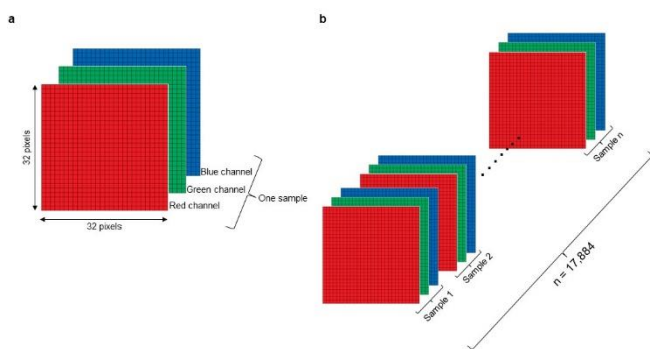


Figure II. Illustration of Numpy 4D arrays. (a) For each sample, 1024 genes are shaped as 32x32 pixels. Due to RGB mapping, 3 layers of color channels were used per sample. (b) This shape was imposed on whole dataset.

### C. CNN Architecture

The CNN architecture shown in Table II was used for training. The architecture includes eight convolution layers, four dropout layers, one global average pooling layer. Each convolution layer consists of 3x3 kernels. ReLU was used as

activation function and to overcome overfitting 0.2 or 0.5 dropout rates were used. Final layer has Sigmoid as activation function.

Table II. CNN Architecture

Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 32, 32, 96)	2,688
dropout (Dropout)	(None, 32, 32, 96)	0
conv2d_1 (Conv2D)	(None, 32, 32, 96)	83,040
conv2d_2 (Conv2D)	(None, 16, 16, 96)	83,040
dropout_1 (Dropout)	(None, 16, 16, 96)	0
conv2d_3 (Conv2D)	(None, 16, 16, 192)	166,080
conv2d_4 (Conv2D)	(None, 16, 16, 192)	331,968
conv2d_5 (Conv2D)	(None, 8, 8, 192)	331,968
dropout_2 (Dropout)	(None, 8, 8, 192)	0
conv2d_6 (Conv2D)	(None, 8, 8, 192)	331,968
activation (Activation)	(None, 8, 8, 192)	0
dropout_3 (Dropout)	(None, 8, 8, 192)	0
conv2d_7 (Conv2D)	(None, 8, 8, 192)	37,056
activation_1 (Activation)	(None, 8, 8, 192)	0
dropout_4 (Dropout)	(None, 8, 8, 192)	0
conv2d_8 (Conv2D)	(None, 8, 8, 2)	386
global_average_pooling2d	(None, 2)	0
activation_2 (Activation)	(None, 2)	0
Total params: 1,368,194		
Trainable params: 1,368,194		
Non-trainable params: 0		

### D. One-Pixel Attack

One pixel attack algorithm was adopted from an earlier study [20] which utilizes "differential evolution" algorithm from SciPy Python library. The attack algorithm picks random locations (x, y) where  $x < 32$  and  $y < 32$  and random RGB colors. Although blue and green values are picked within (0,255) range, the red color was only picked within (0,2) range since gene expression values are mostly below 196,607 corresponding to (2,255,255) RGB value. One pixel attack provides pixel location, the new color value which causes label to change in trained model (from Normal to Tumor or vice versa). Since attack is random, we performed many attacks (10 times to be exact) to the test dataset. The resulting attacks were filtered if the suggested pixel value is within lowest and highest expression range of corresponding gene.

## III. Results

### A. Input Images Obtained by Applying RGB Mapping Method

Since gene expression data have been converted into RGB format, visualizing the expression layout for any sample was possible. In Figure III, sample images for Normal and Tumor samples are presented. Figure III shows 4 sample images from (a) Normal tissue data and (b) Tumor tissue data generated by converting gene expression levels of 1024 selected genes using RGB mapping. The images do not reveal any apparent pattern for naked eye. However, convolutional layers are able to pick regions or patterns formed by neighboring pixels so gene expression data was passed through convolution layers. Please note that gene expression data was converted into RGB format but they are not saved as images before training.

Table III. Comparison model with other studies. SVM; support vector machine, t-SNE; t-distributed stochastic neighbor embedding.

Authors	Expression Preprocessing	Classification	Accuracy	Sensitivity	Specificity	Precision	F-measure
Elbashir et al. [22]	Normalization	CNN	98,76 %	91,43%	100,00%	100,00%	0,955
Danaee et al. [23]	Normalization	Stacked Denoising	94,78 %	94,04%	97,50%	97,20%	
Elbashir et al. [22]	Normalization	AlexNet	96,69 %	96,89%	94,12%	99,54%	0,955
Elbashir et al. [22]	t-SNE	SVM	100,00%	100,00%	51,00%	95,96%	0,97
Proposed method	RGB mapping	CNN	97,73 %	97,66%	97,80%	98,00%	0,975

The training was performed on 32x32x3 3D multidimensional array for each sample.

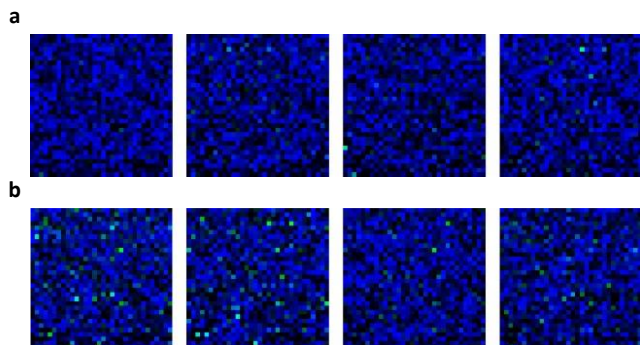


Figure III. Visualization of gene expression data as image.

### B. Model Training

The deep learning architecture shown in Table II has been performed. Evenly distributed normal and tumor samples were trained at a rate of 80:20 using 40 epochs. The accuracy value was determined to be 97.7%. The accuracy and loss plots of the test and training samples are shown in Figure IV.

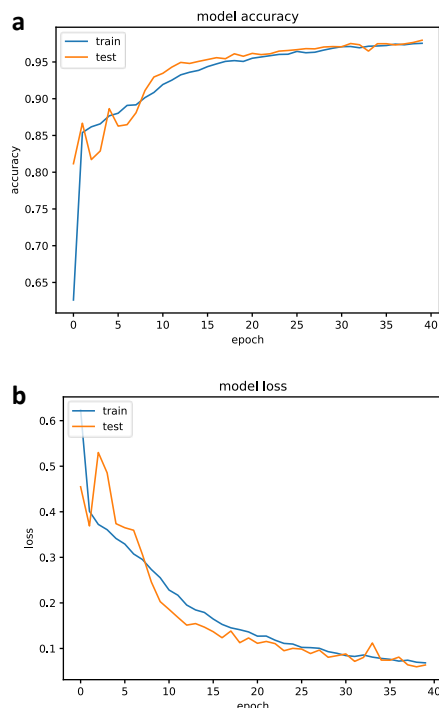


Figure IV. Proposed model (a) model accuracy, and (b) loss

### C. Performance Measurement

Figure V shows the ROC curve of the model. The AUC value of our model is 0.97. There are several different approaches which use gene expression data to classify tumor and normal samples ranging from simpler machine learning approaches to complex deep learning networks. These approaches usually start with pre-processing the gene expression data with an irreversible manipulation and even mapping data points to a different domain. Our method involves minimal and reversible change to gene expression data. The RGB mapping is reversible and does not require normalization or any dimensional reduction techniques. Table III compares our approach with several different approaches both in pre-processing and classification steps.

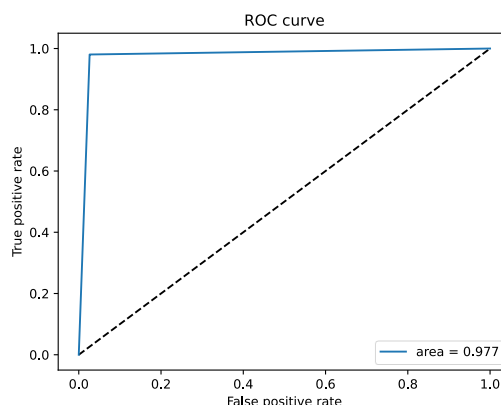


Figure V. The ROC curve of CNN model for tumor and normal classification.

Although Elbashir et al [22] study (Normalization + CNN) has highest accuracy, our approach has better results in overall. Please note that Elbashir et al uses smaller and unbalanced TCGA dataset. Their accuracy starts from 91% and reaches 98.7% and due to dominating number of tumor samples, their model has tendency to pick “tumor” as label, explaining their lowest sensitivity and full precision. In our case, our dataset is balanced (8156 Normal vs. 9750 Tumor) and our accuracy start from 58% and reaches 97.7%.

### D. One-Pixel Attack Results

As a result of the attacks, label changes took place in 240 different samples. When the label changes as a result of the attacks were analyzed, it was seen that the changes on 13 genes caused changes in the decision mechanism of the neural network. The 13 identified genes are shown in Table IV along with their names and Ensembl IDs.

Table IV. Gene List obtained by One-Pixel Attack

Cancer Related Genes According to One-Pixel Attack	<i>Ensembl ID</i>	<i>Gene Name</i>
	ENSG00000163513	TGFBR2
	ENSG00000129250	KIF1C
	ENSG00000215301	DDX3X
	ENSG00000188157	AGRN
	ENSG00000138821	SLC39A8
	ENSG00000124942	AHNAK
	ENSG00000157557	ETS2
	ENSG00000177469	CAVIN1
	ENSG00000123095	BHLHE41
	ENSG00000157514	TSC22D3
	ENSG00000116701	NCF2
	ENSG00000198911	SREBF2
	ENSG00000121691	CAT

When the genes obtained as a result of one pixel attack are examined in the literature, it is seen that the expression changes are associated with the formation of cancer or the survival of the patient. For example, studies have been carried out that the expression changes of the TGFBR2 gene (transforming growth factor beta receptor 2) affect the prognosis in cervical cancer [24] and gastric cancers [25]. Likewise, there are data showing that KIF1C (Kinesin-like protein KIF1C) and SLC39A8 (Solute carrier family 39 member 8) are prognostic marker in renal cancers [26] and that the change in expression level is related to survival time.

When the changes in the estimates of the samples were examined by applying the one pixel attack method, it was seen that only the changes of the AGRN (Agrin) gene occurred by increasing and decreasing the expression level. When the sample images converted to RGB format are examined, the changes in the increased and decreased pixels are seen Figure VI.

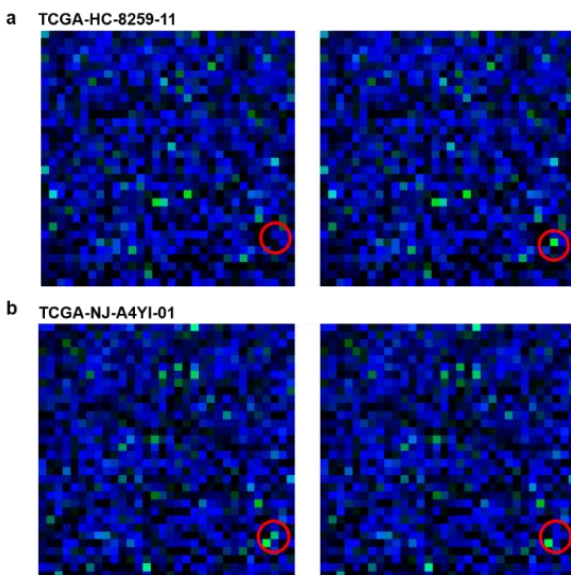


Figure VI. Sample images obtained as a result of the attack.

In Figure VI, the first images show the original images and the second images show the images obtained as a result of the attack. Figure VI (a) shows the original and post attack images of gene expression data with sample ID TCGA-HC-8259-11. The areas marked with a red circle show the changes in the Agrin gene as a result of the attack. The gene expression value has been increased for the TCGA-HC-8259-11 sample. As can be seen in the image, brighter pixels were obtained by increasing the expression value. It has been shown that important clues for the mechanism of cancer can be obtained by performing changes that cannot be distinguished with the naked eye as well as finding effective genes for cancer by attacking the model. In Figure VI (b), the image of the sample with the sample ID TCGA-NJ-A4YI-01 is shown. In this example, the gene expression value for the same gene was decreased. As can be seen in the picture, while the pixel value of the original picture is brighter, it appears darker with decreasing the gene expression value as a result of the attack.

#### IV. Discussion

In this study, a training was carried out through deep learning models and gene expression data of cancer patients using RGB mapping data preprocessing methods. As a result of this training, it has been shown that deep learning methods can distinguish the differences between tumor and normal tissues. The data processing method applied before the model makes it possible to apply a one-pixel attack to the sample images obtained. Identifying genes that are effective for cancer is critical for cancer diagnosis and treatment. For this reason, one pixel attack algorithm was applied to identify genes that may be cancer biomarkers over the obtained training data. When the genes determined by this algorithm were investigated in the literature, it was seen that the expression changes of the genes were effective in cancer progression and survival of patients.

The results of the study have proven that by developing appropriate processing methods for the experimentally obtained biological data, meaningful results can be obtained on the disease without loss of information about the disease. The gene expression data, which are the inputs of the deep learning model, are converted to RGB and applied to the model, allowing the data to be used without any statistical methods (such as normalization) and without loss. In this way, high learning rate and high prediction rate were seen as a result of the training.

All these findings have shown that it can bring a new approach to diseases such as cancer that are difficult to diagnose and require more biomarker genes. With the application of this method, individual results can be obtained. Inter- and intra-tumor heterogeneity characteristics of tumor cells can be determined. It can be used as an approach that makes it possible to make individual cancer analysis by making it easier to find genes that differ from person to person. The results obtained can be strengthened with experimental data to identify new biomarkers for cancer and can be used in personalized diagnostic or therapeutic studies.

#### References

- [1] Esteva, A., Robicquet, A., Ramsundar, B., Kuleshov, V., DePristo, M., Chou, K., Cui, C., Corrado, G., Thrun, S. and Dean,

- J., 2019. A guide to deep learning in healthcare. *Nature medicine*, 25(1), pp.24-29
- [2] Persi, E., Wolf, Y.I., Horn, D., Ruppini, E., Demichelis, F., Gatenby, R.A., Gillies, R.J. and Koonin, E.V., 2020. Mutation–selection balance and compensatory mechanisms in tumour evolution. *Nature Reviews Genetics*, pp.1-12.
- [3] Zuluaga-Gomez, J., Al Masry, Z., Benagoune, K., Meraghni, S. and Zerhouni, N., 2020. A CNN-based methodology for breast cancer diagnosis using thermal images. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, pp.1-15.
- [4] Gour, M., Jain, S. and SunilKumar, T., 2020. Residual learning based CNN for breast cancer histopathological image classification. *International Journal of Imaging Systems and Technology*.
- [5] Swiderska-Chadaj, Z., de Bel, T., Blanchet, L., Baidoshvili, A., Vossen, D., van der Laak, J. and Litjens, G., 2020. Impact of rescanning and normalization on convolutional neural network performance in multi-center, whole-slide classification of prostate cancer. *Scientific Reports*, 10(1), pp.1-14.
- [6] Hartenstein, A., Lübke, F., Baur, A.D., Rudolph, M.M., Furth, C., Brenner, W., Amthauer, H., Hamm, B., Makowski, M. and Penzkofer, T., 2020. Prostate Cancer Nodal Staging: Using Deep Learning to Predict 68 Ga-PSMA-Positivity from CT Imaging Alone. *Scientific Reports*, 10(1), pp.1-11.
- [7] Kanavati, F., Toyokawa, G., Momosaki, S., Rambeau, M., Kozuma, Y., Shoji, F., Yamazaki, K., Takeo, S., Iizuka, O. and Tsuneki, M., 2020. Weakly-supervised learning for lung carcinoma classification using deep learning. *Scientific Reports*, 10(1), pp.1-11.
- [8] Lai, Y.H., Chen, W.N., Hsu, T.C., Lin, C., Tsao, Y. and Wu, S., 2020. overall survival prediction of non-small cell lung cancer by integrating microarray and clinical data with deep learning. *Scientific reports*, 10(1), pp.1-11.
- [9] Jiang, D., Liao, J., Duan, H., Wu, Q., Owen, G. Shu, C., Chen, L., He, Y., Wu, Z., He, D. and Zhang, W., 2020. A machine learning-based prognostic predictor for stage III colon cancer. *Scientific reports*, 10(1), pp.1-9.
- [10] Fontaine, P., Acosta, O., Castelli, J., De Crevoisier, R., Müller, H. and Depeursinge, A., 2020. The importance of feature aggregation in radiomics: a head and neck cancer study. *Scientific Reports*, 10(1), pp.1-11.
- [11] Tschandl, P., Rinner, C., Apalla, Z., Argenziano, G., Codella, N., Halpern, A., Janda, M., Lallas, A., Longo, C., Malvehy, J. and Paoli, J., 2020. Human–computer collaboration for skin cancer recognition. *Nature Medicine*, 26(8), pp.1229-1234.
- [12] Jiao, W., Atwal, G., Polak, P., Karlic, R., Cuppen, E., Danyi, A., De Ridder, J., van Herpen, C., Lolkema, M.P., Steeghs, N. and Getz, G., 2020. A deep learning system accurately classifies primary and metastatic cancers using passenger mutation patterns. *Nature communications*, 11(1), pp.1-12.
- [13] Mencattini, A., Di Giuseppe, D., Comes, M.C., Casti, P., Corsi, F., Bertani, F.R., Ghibelli, L., Businaro, L., Di Natale, C., Parrini, M.C. and Martinelli, E., 2020. Discovering the hidden messages within cell trajectories using a deep learning approach for in vitro evaluation of cancer drug treatments. *Scientific reports*, 10(1), pp.1-11.
- [14] Ramirez, R., Chiu, Y. C., Zhang, S., Ramirez, J., Chen, Y., Huang, Y., & Jin, Y. F., 2021. Prediction and interpretation of cancer survival using graph convolution neural networks. *Methods*.
- [15] Xie, Y., Meng, W. Y., Li, R. Z., Wang, Y. W., Qian, X., Chan, C., ... & Leung, E. L. H., 2021. Early lung cancer diagnostic biomarker discovery by machine learning methods. *Translational oncology*, 14(1), 100907.
- [16] Binder, A., Bockmayr, M., Hägele, M., Wienert, S., Heim, D., Hellweg, K., ... & Klauschen, F. (2021). Morphological and molecular breast cancer profiling through explainable machine learning. *Nature Machine Intelligence*, 1-12.
- [17] Ahn, T., Goo, T., Lee, C. H., Kim, S., Han, K., Park, S., & Park, T., 2018. Deep learning-based identification of cancer or normal tissue using gene expression data. In *2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)* (pp. 1748-1752). IEEE.
- [18] Vivian, J., Rao, A. A., Nothaft, F. A., Ketchum, C., Armstrong, J., Novak, A., ... & Paten, B. (2017). Toil enables reproducible, open source, big biomedical data analyses. *Nature biotechnology*, 35(4), 314-316.
- [19] Rouillard, A. D., Gundersen, G. W., Fernandez, N. F., Wang, Z., Monteiro, C. D., McDermott, M.G., & Ma'ayan, A., 2016. The harmonizome: a collection of processed datasets gathered to serve and mine knowledge about genes and proteins. *Database*, 2016.
- [20] Su, J., Vargas, D. V., & Sakurai, K. (2019). One pixel attack for fooling deep neural networks. *IEEE Transactions on Evolutionary Computation*, 23(5), 828-841.
- [21] O. Dulgerci. (2019). Minimizing with differential evolution. (Visited on 2021-6-18), [Online]. Available: <https://mathematica.stackexchange.com/questions/193009/minimizing-with-differential-evolution>
- [22] Elbashir, M. K., Ezz, M., Mohammed, M., & Saloum, S. S. (2019). Lightweight convolutional neural network for breast cancer classification using RNA-seq gene expression data. *IEEE Access*, 7, 185338-185348.
- [23] Danaee, P., Ghaeini, R., & Hendrix, D. A. (2017). A deep learning approach for cancer detection and relevant gene identification. In *Pacific symposium on biocomputing 2017* (pp. 219-229).
- [24] Yang, H., Zhang, H., Zhong, Y., Wang, Q., Yang, L., Kang, H., ... & Zhou, Y. (2017). Concomitant underexpression of TGFBR2 and overexpression of hTERT are associated with poor prognosis in cervical cancer. *Scientific reports*, 7(1), 1-14.
- [25] Nadauld, L. D., Garcia, S., Natsoulis, G., Bell, J. M., Miotke, L., Hopmans, E. S., ... & Ji, H. P. (2014). Metastatic tumor evolution and organoid modeling implicate TGFBR2 as a cancer driver in diffuse gastric cancer. *Genome biology*, 15(8), 1-18.
- [26] KIF1C, <https://www.proteinatlas.org/ENSG00000129250-KIF1C/pathology/renal+cancer>, 29/05/21.