



Balkan Journal of Electrical & Computer Engineering

An International Peer Reviewed, Referred, Indexed and Open Access Journal

www.bajece.com

Vol :11

No : 1

Year :2023

ISSN : 2147 - 284X



It is abstracted and indexed in, Index Google Scholarship, the PSCR, Cross ref, DOAJ, Research Bible, Indian Open Access Journals (OAJ), Institutional Repositories (IR), J-Gate (Informatics India), Ulrich's, International Society of Universal Research in Sciences, DRJI, EyeSource, Cosmos Impact Factor, Cite Factor, SIS Scientific Indexing Service, IJIF, iijFactor, ULAKBİM-TR Dizin.

General Publication Director & Editor-in-Chief
Musa Yılmaz, University of California Riverside, US

Vice Editor
Hamidreza Nazaripouya, Oklahoma State University, US

Scientific Committee
Abhishek Shukla (India)
Abraham Lomi (Indonesia)
Aleksandar Georgiev (Bulgaria)
Arunas Lipnickas (Lithuania)
Audrius Senulis (Lithuania)
Belle R. Upadhyaya (USA)
Brijender Kahanwal (India)
Chandar Kumar Chanda (India)
Daniela Dzhonova-Atanasova (Bulgaria)
Deris Stiawan (Indonesia)
Emel Onal (Turkey)
Emine Ayaz (Turkey)
Enver Hatimi (Kosovo)
Ferhat Sahin (USA)
Gursel Alici (Australia)
Hakan Temeltaş (Turkey)
Ibrahim Akduman (Turkey)
Jan Izykowski (Poland)
Javier Bilbao Landatxe (Spain)
Jelena Dikun (Lithuania)
Karol Kyslan (Slovakia)
Kunihiko Nabeshima (Japan)
Lambros Ekonomou (Greece)
Lazhar Rahmani (Algerie)
Marcel Istrate (Romania)
Marija Eidukeviciute (Lithuania)
Milena Lazarova (Bulgaria)
Muhammad Hadi (Australia)
Muhamed Turkanović (Slovenia)
Mourad Houabes (Algerie)
Murari Mohan Saha (Sweden)
Nick Papanikolaou (Greece)
Okyay Kaynak (Turkey)
Osman Nuri Ucan (Turkey)
Ozgur E. Mustecaplioglu (Turkey)
Padmanaban Sanjeevikumar (India)
Ramazan Caglar (Turkey)
Rumen Popov (Bulgaria)
Tarek Bouktir (Algeria)
Sead Berberovic (Croatia)
Seta Bogosyan (USA)
Savvas G. Vassiliadis (Greece)
Suwarno (Indonesia)
Tulay Adali (USA)
Yogeshwarsing Calleecharan (Mauritius)
YangQuan Chen (USA)
Youcef Soufi (Algeria)

Aim & Scope

The journal publishes original papers in the extensive field of Electrical-Electronics and Computer engineering. It accepts contributions which are fundamental for the development of electrical engineering, computer engineering and its applications, including overlaps to physics. Manuscripts on both theoretical and experimental work are welcome. Review articles and letters to the editors are also included.

Application areas include (but are not limited to): Electrical & Electronics Engineering, Computer Engineering, Software Engineering, Biomedical Engineering, Electrical Power Engineering, Control Engineering, Signal and Image Processing, Communications & Networking, Sensors, Actuators, Remote Sensing, Consumer Electronics, Fiber-Optics, Radar and Sonar Systems, Artificial Intelligence and its applications, Expert Systems, Medical Imaging, Biomedical Analysis and its applications, Computer Vision, Pattern Recognition, Robotics, Industrial Automation.



ISSN: 2147- 284X
Vol: 11
No : 1
Year: January 2023

CONTENTS

Semiha Koşu, Serdar Özgür Ata; NOMA-Enabled Cooperative V2V Communications with Fixed-Gain AF Relaying,	1-12
Muhammet Fatih Aslan; Comparative Analysis of CNN Models and Bayesian Optimization-Based Machine Learning Algorithms in Leaf Type Classification,	13-24
Sadiq Iqbal, Jihad M. Hamamreh; Precoded Universal MIMO Superposition Transmission for Achieving Optimal Coverage and High Throughput in 6G and Beyond Networks,	25-34
Seda Üstün Ercan, Mohammed Sufyan Mohammed; IoT and XBee Based Central Car Parking Management System,	35-41
Ali Burak Öncül; LSTM-GRU Based Deep Learning Model with Word2Vec for Transcription Factors in Primates,	42-49
Sevilay Tüfenkçi, Barış Baykant Alagöz, Celalettin Yeroğlu, Bilal Şenol; A Software Realization of Disturbance Rejection Optimal FOPID Controller Design Methodology by Using Soft Computing Techniques,	50-60
Tolga Ulutaş, Osman Avcı, Engin Can Akar, Barış Köksal, Yılmaz Kalkan; Simple Design and Implementation of Two-Way Communication System through UAV,	61-70
Nur Hüseyin Kaplan; Single Image Dehazing based on Additive Wavelet Transform,	71-77
Emrah Aydemir, Raghad Tohmas Esfandiyar Alalawi; Classification Of Hand Images by Person, Age and Gender with The Median Robust Extended Local Binary Model,	78-87
Hasan Huseyin Coban; Analyzing the Societal Cost of Electric Roads Compared to Batteries and Oil for All Forms of Road Transport,	88-99
İsa Ataş; Performance Evaluation of Jaccard-Dice Coefficient on Building Segmentation from High Resolution Satellite Images,....	100-106

BALKAN JOURNAL OF ELECTRICAL & COMPUTER ENGINEERING

(An International Peer Reviewed, Indexed and Open Access Journal)

Contact

Batman University
Department of Electrical-Electronics Engineering
Bati Raman Campus Batman-Turkey

Web: <http://dergipark.gov.tr/bajece>
<https://www.bajece.com>
e-mail: bajece@hotmail.com

NOMA-enabled Cooperative V2V Communications with Fixed-Gain AF Relaying

Semiha Koşu and Serdar Özgür Ata

Abstract—By virtue of its improving bandwidth efficiency along with user fairness, non-orthogonal multiple access (NOMA) technique is considered a promising method for next-generation wireless communication systems. Since fading effect of wireless channels in vehicle-to-vehicle (V2V) communication systems are more severe than those in traditional systems, in this study, we employ the power-domain downlink NOMA technique in cooperative V2V communication systems to enhance data transmission capacity and network efficiency. In the proposed system, the base station communicates with two vehicular nodes, namely near and far users, through a relay vehicle employing the fixed-gain amplify-and-forward scheme. In real-life scenarios, the relay and the users move in high velocities; hence, the corresponding fading channels between these nodes are exposed as having double-Rayleigh fading characteristic in which the fading coefficient of a wireless channel is modeled as the product of two Rayleigh distributed random variables. To analyze the system performance, we first investigate the outage probability and derive its exact closed-form expressions for the near and far users. Then, we make the exact ergodic capacity analysis and obtain the closed-form solution for the near user. Furthermore, outage and ergodic performances of the NOMA-enabled system are compared to the simulation results of the traditional orthogonal multiple access approach. We also give analytic and numerical results to evaluate the performance of the proposed system and show the consistency of Monte-Carlo simulations with analytical derivations. It is observed that even with the small power allocation, both performances of the near user mostly outperform the far user.

Index Terms—Amplify-and-forward, cooperative communications, double Rayleigh, ergodic capacity, fixed-gain, non-orthogonal multiple access, outage probability, vehicular networks.

I. INTRODUCTION

IN wireless communication systems, multiple access methods are used in order to allocate limited resources effectively to many users [1]. Well-known orthogonal multiple access (OMA) techniques used in current cellular communication systems, such as time division multiple access, frequency division multiple access or code division multiple access, cannot fulfill the higher spectral efficiency and enhanced user fairness requirements of 5G and beyond technologies. Recently, non-orthogonal multiple access (NOMA) has been proposed to be an outperforming multiple access technique which offers higher spectral efficiency in comparison with OMA techniques

by allowing multiple users to simultaneously transmit their signal with different power levels in the same resource block (i.e., frequency, time, code) [2]. In addition to enhanced spectrum usage efficiency, NOMA can ensure high fairness among multiple users through optimized power allocation schemes, in which users suffering poor channel conditions, namely weak/far users, receive their signals with a higher power while users experiencing better channel conditions, namely strong/near users, receive lower power information symbols. Hence, the strong users achieving high signal-to-interference-plus-noise ratio (SINR) will be able to decode their own signal by using successive interference cancellation (SIC) methods while the weak ones having low SINR may decode their own signals by treating the strong users' signals as interference.

Vehicle-to-vehicle (V2V) communications has attracted increasing interest as an important application area of 5G technology with the opportunities it offers such as in-vehicle internet access and autonomous driving, as well as intersection management and preventing densities by regulating the traffic flow, collision avoidance and rapid response to accidents, roadway reservation and increasing road safety [3]–[8]. Compared to cellular communication systems, the design of V2V communication systems has its own challenges. One of these is that the fading effect in wireless communication channels is worse compared to the fading effect encountered in cellular communication systems due to factors such as high speeds of vehicles, low vehicle-antenna heights, and high number of scattering obstacles [9]. Many studies in the literature have shown that the fading effect in V2V communication channels can be modeled as the product of several random variables. These types of channels are named cascaded fading channels. Again in these studies, it has been revealed by field measurements that a suitable channel model for modeling the V2V communication channels is the double Rayleigh fading channel model in which the fading gain of the wireless channel can be expressed as the product of two independent Rayleigh distributed random variables [9]–[12]. However, there are very few studies in the literature focusing on V2V communication systems over cascaded fading channels compared to studies on traditional cellular communication systems. This is mainly because the closed-form solutions for the performance analysis of these channels cannot be obtained most of the time due to intractable integrals resulting from special functions included in the cumulative distribution function (CDF) and probability density function (PDF) expressions of SINR.

Semiha Koşu is with the Department of Informatics Institute, Istanbul Technical University, 34469 Istanbul, TURKEY e-mail: kosu18@itu.edu.tr

Serdar Özgür Ata is with the Informatics and Information Security Research Center, TUBITAK-BILGEM, Kocaeli, TURKEY e-mail: serdar.ata@tubitak.gov.tr

Manuscript received Mar 21, 2022; accepted Dec 30, 2022. DOI: 10.17694/bajece.1090937

A. Related Works

As an effective performance-enhancer technique, cooperative communications has been studied extensively over two decades. Amplify-and-forward (AF) and decode-and-forward (DF) relaying methods [1] were proposed to achieve better performance in many V2V communication scenarios where line-of-sight (LOS) is not available between source and destination nodes due to reasons such as vehicles with low-antenna heights in urban areas or communication through very long distances in rural areas. In the literature, cooperative V2V systems over cascaded fading channels are studied in [13]–[18]. In [13], outage probability and achievable diversity order are examined for DF relaying in V2V communication systems by applying the best relay selection scheme over cascaded Rayleigh fading channels. In [14], by using the moment generation function approach, higher-order moments of the signal-to-noise ratio (SNR), average symbol error probability, and ergodic capacity are derived for the bi-directional AF transmission system over cascaded Nakagami- m channels. In [15], the probability of detection is investigated in an inter-vehicular cooperative cognitive radio system using the maximum ratio combining (MRC) diversity technique when the channels are exposed to cascaded Rayleigh fading. In [16], the minimized total error rate and the probability of detection are analyzed over mixture Gamma distributed fading channels for a cooperative spectrum sensing system employing different diversity combining techniques with multiple antenna nodes. In [17], the authors studied a cooperative V2V interference-limited communication system with DF relaying operating over double-Nakagami fading channels. Exact and asymptotic expressions for outage and average symbol error probabilities are obtained for the cases of best relay selection and threshold-based relay selection schemes. The outage performance of V2V networks employing full-duplex (FD) and half-duplex (HD) relaying schemes is investigated in [18] and the upper bounds for the outage probability are derived for AF and DF relaying over double Rayleigh fading channels.

In recent years, in order to fulfill robustness and reliability requirements of next generation technologies, NOMA and cooperative communication techniques are addressed together in system design [19]–[29]. In [19], asymptotic outage probability and ergodic sum-rate analyses are derived for a NOMA-inspired AF relaying system where the source and the relay simultaneously transmit two symbols in a message. Additionally, in [19], the authors propose a mixed MRC-SIC scheme for received superimposed signals at the destination. In [20], the authors applied the NOMA technique to an AF relaying system to maximize the achievable secrecy sum rate by jointly designing power allocation at the source and cooperative beam-forming at the relay, under achievable rate constraints at the weak users and transmit power constraints at the source and the relay. [21] investigates the performance of AF relaying scheme using NOMA, where two sources communicate with their corresponding destinations over the same frequency simultaneously through a shared AF relay, and closed-form approximation for the outage probability and an upper bound to the ergodic sum capacity are derived.

[22] proposes a self-energy recycling FD cooperative NOMA system, where a nearby user can be employed as a DF or AF relay with self-energy recycling protocol to assist a far user, and new expressions of exact and asymptotic outage probabilities for two users are derived. In [23], downlink cooperative multiple-input single-output wireless sensor networks with NOMA over Nakagami- m fading are studied. Novel antenna-relay-destination selection schemes are proposed for DF and AF relaying strategies, and then, closed-form expressions of outage probability at the selected sensor nodes are derived. [24] investigates the performance of a virtual FD relaying technique in cooperative NOMA-based systems in which two HD AF relays imitate the operation of FD relaying, and the closed-form expression of outage probability is derived. In [25], a power-domain NOMA system with partial statistical channel state information (CSI) is studied and ergodic sum rate and outage probability are obtained for both DF and AF relaying strategies over Nakagami- m fading channels. In [26], the authors obtained closed-form expressions of the outage probabilities for a two-stage relay selection strategy for DF and AF relaying protocols employed in NOMA networks where one base station (BS) communicates with two mobile users with the aid of multiple relays. The outage performance of a cooperative NOMA-based network that applies AF relaying protocol is studied in [27], and approximate outage probability and its asymptotic behaviors are obtained. In [28], performance of the fixed-gain AF relaying with NOMA is investigated over Nakagami- m fading channels for two different scenarios where BS intends to communicate with multiple users through the assistance of AF relaying with and without LOS. A hybrid AF and DF with NOMA transmission scheme is proposed for a cellular system with multiple relays in [29]. It is shown that the proposed scheme can achieve larger sum channel capacity and larger average system throughput compared with the traditional schemes.

Furthermore, different system models for NOMA-enabled vehicular communication infrastructures are investigated throughout several works in the literature [30]–[35]. In [30], the exact expression of the average achievable rate of a NOMA-based cooperative relaying system over Rician fading channels is obtained. Outage probability expressions are derived for near and far user nodes whether they are on the roads (vehicle, cyclist, pedestrian) or outside the roads (BS, road side unit) in [31]. The secrecy outage probability performance of a NOMA-based vehicular communication system is analyzed and closed-form analytical expressions for the FD/HD relaying protocols over Nakagami- m fading channels are derived in [32]. The outage performance analysis of cooperative vehicular networks is made and the closed-form outage probability expressions are obtained in [33] when the MRC technique is applied at the relay. In [34], a novel cooperative transmission scheme employing a signal superposition technique is proposed and a closed-form expression for the link reliability of the proposed scheme is derived. It is shown that high communication reliability may be achieved with this technique since vehicle user equipments (VUEs) superpose other VUEs' signals and then re-transmit other users' V2V packets. [35] analyzes the security performance

of the FD-NOMA based V2V system from the physical layer perspective, and then derives analytical results of the ergodic secrecy capacity. In this study, the authors also propose a secrecy sum rate optimization scheme utilizing the instantaneous CSI. In [36], two half-duplex/full-duplex relay-assisted NOMA-based scheme is investigated to solve optimum power allocation problem among the users to improve the quality-of-service with the poor channel conditions over independent Rayleigh distributed channels. The BER and ergodic capacity performance of NOMA-SM scheme in V2V Massive MIMO system has been investigated in [37]. Moreover, in [37], power allocation optimization scheme is studied over a spatio-temporally correlated Rician channel.

It should be noted that the channel models in all the abovementioned studies were selected from among channel models valid for cellular communication systems, and NOMA-based cooperative V2V communication systems over cascaded fading channels have not been studied in the literature yet. Additionally, performance of non-cooperative vehicular systems with NOMA over cascaded fading channels is investigated in [38]–[40]. In [38], the overall outage probability of two user-assisted NOMA-based V2V systems is studied over double Rayleigh fading channels. Joint power allocation optimization and the probability of successful decoding for caching-aided 2-user NOMA-based system is investigated for double Nakagami- m fading channels in [39]. In [40], a vehicle clustering method is proposed and maximum achievable rate between clusters is studied over double Rayleigh fading channels.

B. Motivation and Contribution

In V2V communication systems, fading channels may be much more hostile due to reasons such as high velocities of user nodes, low antenna-heights, and mostly unavailable LOS between transmitter and receiver. Therefore, to achieve robust communication at high data rates over far distances, some performance enhancement methods must be considered in the system design. Additionally, it has been shown that the fading coefficients of vehicular channels may be modeled as the multiplication of a number of random variables, and several studies in the literature show that the double Rayleigh fading model can adequately model this type of channels [10]–[12], [38], [40]. Additionally, in order to avoid the mathematical complexity, some papers in the literature consider the fading effects at the V2V links as Rayleigh fading or other fading models which are better than Rayleigh fading [41]–[43]. On the other hand, in our proposed system, V2V channels are assigned to be subject to the double Rayleigh fading which is the most accurately fitting fading model for the considered environment. At this point, we emphasize that for the first time in the literature, we propose to exploit the benefits of NOMA and fixed-gain AF relayed cooperative communication techniques together for V2V systems, and then investigate the performance of the proposed system in terms of outage probability and ergodic capacity metrics. Hence the main contributions of this paper can be summarized as follows:

- For the first time in the literature, a power-domain downlink NOMA-enabled cooperative vehicular communication

system employing fixed-gain AF relaying scheme is proposed to cope with the severe fading conditions over double Rayleigh fading channels.

- We examine the performance of the proposed system by deriving the exact analytical expressions of outage probabilities in closed-form for near and far users over double Rayleigh fading channels.
- We obtain closed-form analytical expressions of the ergodic capacity to provide more insight into the vehicular communication systems' performance over double Rayleigh fading channels.
- We present the analytic and numerical results with respect to different system parameters such as transmitting power, distance between nodes, and the power allocation metrics to evaluate the performance of the vehicular communication system.
- Finally, we compared the NOMA system performance with the traditional OMA technique considering the outage probability and ergodic capacity analyses.

The remainder of the manuscript is organized as follows: Section II details the proposed system and the channel model. Section III investigates the outage probabilities at the near and far users. In Section IV, ergodic capacity for both users and the overall system are examined. In Section V, analytic and numerical results are provided to evaluate the system's performance and to verify the analytic result via Monte-Carlo simulations. Section VI concludes the paper.

Throughout this paper, $F_X(\cdot)$ and $f_X(\cdot)$ denote the CDF and the PDF of a random variable X , respectively, while $E[\cdot]$ symbolizes the expectation operator.

II. SYSTEM MODEL

One of the major distinctions between vehicular and conventional cellular communication systems is the antenna elevation of the nodes. In the proposed system, all vehicle nodes are assumed to be equipped with the antennas at low elevations and to be in motion, which results in local scattering around the vehicles causing delay and attenuation for the received signal. In this study, we consider a vehicular communication system with power-domain downlink NOMA which consists of a BS, a relay (R) and two paired users, i.e., U_1 (strong user) and U_2 (weak user), as illustrated in Fig. 1. As a realistic scenario for the V2V communications in urban areas in which the LOS between BS and the end-users may not be available most of the time due to the high buildings around the vehicles, R is used as an intermediate relaying node to help BS communicate with the users. We assume that all nodes have a single antenna and operate in HD mode. Further, h_r , h_1 and h_2 are the fading coefficients of the wireless channels between BS–R, R– U_1 , and R– U_2 , respectively. Also the distances between BS–R, R– U_1 , and R– U_2 are denoting by d_{sr} , d_1 , and d_2 , respectively, and they are normalized with respect to the total distance between BS– U_2 , which is $d_{sr} + d_2$. Since R, U_1 and U_2 are nodes moving at high speeds, it is assumed that the corresponding channel coefficients h_1 and h_2 are double Rayleigh distributed while the channel coefficient h_r for BS→R has Rayleigh distribution due to the

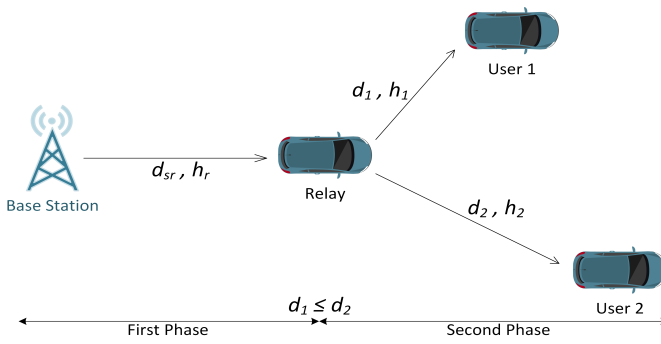


Fig. 1. NOMA-enabled cooperative vehicular communication system employing fixed-gain AF relaying.

fixed BS. Here, h_1 and h_2 can be expressed by a product of two single Rayleigh distributions, which in turn results in the envelope of the channel having a double Rayleigh fading distribution. Therefore, considering the double Rayleigh fading channel model, which is worse than the Rayleigh fading, for the vehicular communication systems is accurately reasonable. Furthermore, without loss of generality, $d_1 \leq d_2$ and the channel gains of the users are ordered as $|h_1|^2 \geq |h_2|^2$ and $E[|h_i|^2] = \Omega_i$ for $i \in \{1, 2\}$.

In the system, information transmission occurs in two phases. During the first phase, BS broadcasts superimposed signals, s_1 for U_1 and s_2 for U_2 , to the node R with a total transmit power of P_t . Hence, the signal transmitted from the BS using the downlink NOMA technique is [44]

$$x = \sqrt{P_1}s_1 + \sqrt{P_2}s_2, \quad (1)$$

where P_1 and P_2 are the allocated powers for U_1 and U_2 , respectively, determined as $P_1 = \alpha_1 P_t$ and $P_2 = \alpha_2 P_t$. Here α_i is the power allocation coefficient and $\alpha_1 + \alpha_2 = 1$. By following the principle of NOMA, BS allocates more power to the weak user's signal than the strong user's signal so that $\alpha_1 < \alpha_2$. Hence, the signal received at R is

$$y_r = h_r x + n_r = h_r(\sqrt{P_1}s_1 + \sqrt{P_2}s_2) + n_r \quad (2)$$

where $n_r \sim \mathcal{CN}(0, N_0)$ is the additive white Gaussian noise (AWGN) at R with power N_0 . Then, in the second transmission phase, R applies the fixed-gain AF relaying technique and forwards the scaled signal to the users. Therefore, the signal received at the i^{th} user becomes

$$\begin{aligned} y_i &= G h_i y_r + n_i \\ &= \underbrace{G h_i h_r (\sqrt{P_1}s_1 + \sqrt{P_2}s_2)}_{\text{Signal broadcast from BS}} + \underbrace{G h_i n_r + n_i}_{\text{Noise}} \end{aligned} \quad (3)$$

where $G = \sqrt{\frac{P_r}{P_t \Omega_r + N_0}}$ is the power normalization factor and $n_i \sim \mathcal{CN}(0, N_0)$ is the AWGN for the i^{th} user.

Due to the higher power level assigned to U_2 's symbols at BS, assuming perfect SIC is performed at the receiver, U_1 first decodes s_2 from the received signal y_1 . Thus, the SINR related to the processing of s_2 is given as

$$\gamma_{1,2} = \frac{G^2 |h_1|^2 |h_r|^2 P_2}{G^2 |h_1|^2 |h_r|^2 P_1 + G^2 |h_1|^2 N_0 + N_0}. \quad (4)$$

After successfully decoding s_2 , U_1 subtracts the corresponding interference term from the received signal and obtains its own symbol s_1 without any interference from the other user. Therefore, the obtained SNR corresponding to the decoding process of s_1 becomes

$$\gamma_1 = \frac{G^2 |h_1|^2 |h_r|^2 P_1}{G^2 |h_1|^2 N_0 + N_0}. \quad (5)$$

In parallel, since s_1 has been transmitted with a small portion of P_t from BS, U_2 can decode its own symbol s_2 directly from the received signal y_2 without employing SIC by treating the corresponding term of s_1 as interference. Hence the SINR achieved by U_2 during the decoding process of s_2 may be expressed as

$$\gamma_2 = \frac{G^2 |h_2|^2 |h_r|^2 P_2}{G^2 |h_2|^2 |h_r|^2 P_1 + G^2 |h_2|^2 N_0 + N_0}. \quad (6)$$

In the following sections, we investigate the performance of the proposed system in terms of outage probabilities and achievable ergodic capacities for both U_1 and U_2 .

III. OUTAGE PROBABILITY ANALYSIS

A. Exact Outage Probability for the Near User

In the proposed system, U_1 falls into outage regime when s_1 or s_2 cannot be decoded successfully by U_1 . In other words, the outage regime might happen if γ_1 or $\gamma_{1,2}$ goes below a certain threshold value and can be probabilistically expressed as

$$P_{out}^1 = 1 - \mathcal{P}\{\gamma_{1,2} > \gamma_{th2}, \gamma_1 > \gamma_{th1}\} \quad (7)$$

where $\gamma_{th1} = 2^{2\mathcal{R}_1} - 1$ and $\gamma_{th2} = 2^{2\mathcal{R}_2} - 1$ are the predefined threshold values for U_1 and U_2 , respectively. Here, \mathcal{R}_1 and \mathcal{R}_2 are the targeted data rates at U_1 and U_2 , respectively. By substituting (4) and (5) into (7), the outage probability is obtained and simplified as

$$\begin{aligned} P_{out}^1 &= 1 - \mathcal{P}\left\{|h_r|^2 > \tau, |h_1|^2 > \left(\frac{\tau/G^2}{|h_r|^2 - \tau}\right)\right\} \\ &\times \mathcal{P}\left\{|h_r|^2 > v, |h_1|^2 > \left(\frac{v/G^2}{|h_r|^2 - v}\right)\right\} \end{aligned} \quad (8)$$

where $\tau \triangleq \varrho_2/(P_2 - P_1\gamma_{th2})$, $v \triangleq \varrho_1/P_1$, $\varrho_1 \triangleq N_0\gamma_{th1}$ and $\varrho_2 \triangleq N_0\gamma_{th2}$. Here, it is assumed that $\gamma_{th2} < (P_2/P_1)$. Please note that, unless otherwise stated, P_{out}^1 is always one. Then, by defining $\phi \triangleq \max(\tau, v)$, (8) is expressed in integral form as

$$\begin{aligned} P_{out}^1 &= 1 - \mathcal{P}\left\{|h_r|^2 > \phi, |h_1|^2 > \frac{\phi/G^2}{|h_r|^2 - \phi}\right\} \\ &= \int_0^\phi f_{|h_r|^2}(y) dy + \int_\phi^\infty f_{|h_r|^2}(y) F_{|h_1|^2}\left(\frac{\phi/G^2}{y - \phi}\right) dy. \end{aligned} \quad (9)$$

Since h_r is exposed to Rayleigh fading, the PDF of $|h_r|^2$ is

$$f_{|h_r|^2}(\gamma) = \frac{1}{\gamma_r} e^{-\frac{\gamma}{\gamma_r}} \quad (10)$$

where $\bar{\gamma}_r = \Omega_r$ [1]. Furthermore, for double Rayleigh fading channels, the CDF of the unordered channel gain $|\tilde{h}|^2$ is [38]

$$F_{|\tilde{h}|^2}(\gamma) = 1 - 2\sqrt{\frac{\gamma}{\bar{\gamma}}} K_1\left(2\sqrt{\frac{\gamma}{\bar{\gamma}}}\right) \quad (11)$$

where $K_v(\cdot)$ is the v^{th} order modified Bessel function of the second kind [45] and $\bar{\gamma} = E[|\tilde{h}|^2]$. Consequently, with the aid of order statistic [46], the CDF of the ordered channel gain in double Rayleigh fading channels becomes

$$F_{|h|^2}(\gamma) = \sum_{k=0}^{M-l} \frac{(-1)^k M!}{(l+k)(M-l-k)! (l-1)! k!} [F_{|\tilde{h}|^2}(\gamma)]^{l+k} \quad (12)$$

where M is the number of users and l is the order index. For the proposed system, $M = 2$ and $l = 2$. Therefore, by substituting (11) in (12), CDF of the ordered channel gain $|h_1|^2$ becomes

$$F_{|h_1|^2}(\gamma) = [F_{|\tilde{h}_1|^2}(\gamma)]^2 = 1 - \omega_1(\gamma, \bar{\gamma}_1) + \omega_2(\gamma, \bar{\gamma}_1) \quad (13)$$

where $\bar{\gamma}_1 = \Omega_1$ while $\omega_1(\cdot, \cdot)$ and $\omega_2(\cdot, \cdot)$ are defined as

$$\omega_1(a, b) \triangleq 4\sqrt{\frac{a}{b}} K_1\left(2\sqrt{\frac{a}{b}}\right) \quad (14a)$$

and

$$\omega_2(a, b) \triangleq \frac{4a}{b} \left[K_1\left(2\sqrt{\frac{a}{b}}\right) \right]^2 \quad (14b)$$

where a is the parameter and $b \in \{\bar{\gamma}_1, \bar{\gamma}_2\}$, respectively. Thus, by using (10), (13), (14a) and (14b), P_{out}^1 may be expressed as

$$P_{out}^1 = 1 - I_1 + I_2 \quad (15)$$

where

$$I_1 = \int_{\phi}^{\infty} f_{|h_r|^2}(\gamma) \omega_1\left(\frac{\phi/G^2}{\gamma - \phi}, \bar{\gamma}_1\right) d\gamma \quad (16)$$

and

$$I_2 = \int_{\phi}^{\infty} f_{|h_r|^2}(\gamma) \omega_2\left(\frac{\phi/G^2}{\gamma - \phi}, \bar{\gamma}_1\right) d\gamma. \quad (17)$$

Substituting (10) and the explicit expression of ω_1 in (16), and by applying $\gamma - \phi = t$ transformation, I_1 is becomes as

$$I_1 = 4\sqrt{\frac{\phi/G^2}{\bar{\gamma}_1}} \kappa_1 \int_0^{\infty} t^{-\frac{1}{2}} e^{-\frac{t}{\bar{\gamma}_r}} K_1\left(2\sqrt{\frac{1}{\bar{\gamma}_1} \frac{\phi/G^2}{t}}\right) dt \quad (18)$$

where $\kappa_1 \triangleq \frac{1}{\bar{\gamma}_r} e^{-\frac{\phi}{\bar{\gamma}_r}}$. After that, by using [47, 14], [45, 9.31-1], [45, 9.31-2], [45, 9.31-5] and [45, 7.813-1], the closed-form solution of (18) is obtained as

$$I_1 = 2e^{-\frac{\phi}{\bar{\gamma}_r}} G_{4,1}^{1,3}\left(\frac{\bar{\gamma}_1 \bar{\gamma}_r}{\phi/G^2} \middle| \begin{matrix} 0, 0, 1, -\frac{1}{2} \\ -\frac{1}{2} \end{matrix}\right) \quad (19)$$

where $G_{p,q}^{m,n}(\cdot|\cdot)$ denotes Meijer's G-function [45]. By substituting (10) and (14b) in (17), with the help of $\gamma - \phi = t$ transformation and after some algebraic simplifications, I_2 is written as

$$I_2 = \kappa_1 \int_0^{\infty} \frac{4\phi/G^2}{\bar{\gamma}_1 t} e^{-\frac{t}{\bar{\gamma}_r}} \left[K_1\left(2\sqrt{\frac{1}{\bar{\gamma}_1} \frac{\phi/G^2}{t}}\right) \right]^2 dt. \quad (20)$$

Furthermore, first applying [47, 11] and [48, 03.04.26.0014.01] and then using [45, 9.31-5], [45, 9.31-1] and [45, 9.31-2], and finally with the help of [47, 21], (20) is solved and the closed-form solution for I_2 is obtained as

$$I_2 = \frac{\sqrt{\pi}}{2} e^{-\frac{\phi}{\bar{\gamma}_r}} G_{6,3}^{2,4}\left(\frac{\bar{\gamma}_1 \bar{\gamma}_r}{4\phi/G^2} \middle| \begin{matrix} 0, 1, -1, 0, -1, -1 \\ -1, -1, -\frac{1}{2} \end{matrix}\right). \quad (21)$$

Finally, by substituting (19) and (21) in (15), the exact outage probability expression for the near user U_1 is obtained in closed-form.

B. Exact Outage Probability for the Far User

From the perspective of the far user, the outage regime occurs when U_2 can not decode its own symbol s_2 successfully due to receiving lower SNR value compared to the threshold value. As for its probability, P_{out}^2 is given as

$$P_{out}^2 = \mathcal{P}\{\gamma_2 \leq \gamma_{th_2}\}. \quad (22)$$

By substituting (6) in (22), P_{out}^2 takes the form

$$P_{out}^2 = \mathcal{P}\left\{ \frac{G^2 |h_2|^2 |h_r|^2 P_2}{G^2 |h_2|^2 |h_r|^2 P_1 + G^2 |h_2|^2 N_0 + N_0} \leq \gamma_{th_2} \right\} = \mathcal{P}\left\{ |h_2|^2 G^2 (|h_r|^2 \lambda - N_0 \gamma_{th_2}) \leq N_0 \gamma_{th_2} \right\} \quad (23)$$

where $\lambda = P_2 - P_1 \gamma_{th_2}$. From (23), it is seen that the outage probability at U_2 can be written as

$$P_{out}^2 = \mathcal{P}\{|h_r|^2 < \tau\} + \mathcal{P}\left\{|h_r|^2 > \tau, |h_2|^2 < \frac{\tau/G^2}{|h_r|^2 - \tau}\right\}. \quad (24)$$

According to (24), the condition $\gamma_{th_2} < P_2/P_1$ must be satisfied system-wide to keep the far user from being exposed to unavoidable outage condition. Further, by expressing (24) in integral form, it is written as

$$P_{out}^2 = \int_0^{\tau} f_{|h_r|^2}(y) dy + \int_{\tau}^{\infty} f_{|h_r|^2}(y) F_{|h_2|^2}\left(\frac{\tau/G^2}{|h_r|^2 - \tau}\right) dy. \quad (25)$$

By using (12) with $M = 2$ and $l = 1$ for U_2 , the CDF of the ordered channel gain $|h_2|^2$ for double Rayleigh fading channels is obtained as

$$F_{|h_2|^2}(\gamma) = 2F_{|\tilde{h}_2|^2}(\gamma) - [F_{|\tilde{h}_2|^2}(\gamma)]^2. \quad (26)$$

Then, by defining $\chi \triangleq 2\sqrt{\frac{\gamma}{\bar{\gamma}_2}} K_1\left(2\sqrt{\frac{\gamma}{\bar{\gamma}_2}}\right)$ where $\bar{\gamma}_2 = \Omega_2$ and substituting (11) in (26), (26) can be written as

$$F_{|h_2|^2}(\gamma) = 2(1 - \chi) - (1 - \chi)^2 = 1 - \chi^2 = 1 - \omega_2(\gamma, \bar{\gamma}_2) \quad (27)$$

Additionally, by substituting (27) into (25), P_{out}^2 is written as

$$P_{out}^2 = 1 - I_3 \quad (28)$$

where

$$I_3 = \int_{\tau}^{\infty} f_{|h_r|^2}(\gamma) \omega_2\left(\frac{\tau/G^2}{\gamma - \tau}, \bar{\gamma}_2\right) d\gamma = \int_{\tau}^{\infty} \frac{4}{\bar{\gamma}_2} \frac{\tau/G^2}{\gamma - \tau} \left[K_1\left(2\sqrt{\frac{1}{\bar{\gamma}_2} \frac{\tau/G^2}{\gamma - \tau}}\right) \right]^2 \frac{1}{\bar{\gamma}_r} e^{-\frac{\gamma}{\bar{\gamma}_r}} d\gamma \quad (29)$$

which is similar to the expression in (20) with the replacement of the parameter ϕ with τ . Therefore, by applying $\gamma - \tau = t$ transformation and first applying [47, 11] and [48, 03.04.26.0014.01], then using [45, 9.31-5], [45, 9.31-1] and [45, 9.31-2], and finally with the help of [47, 21], (29) is solved and the closed-form solution for P_{out}^2 is obtained as

$$P_{out}^2 = 1 - \frac{\sqrt{\pi}}{2} e^{-\frac{\tau}{\gamma_r}} G_{5,2}^{1,4} \left(\frac{\gamma_2 \gamma_r}{4\tau/G^2} \middle| \begin{matrix} 0,0,1,-1,0 \\ 0,-\frac{1}{2} \end{matrix} \right). \quad (30)$$

C. System Outage Probability

In this section, the overall system outage probability is defined as the probability of at least one user being exposed to outage through the system. Therefore, the overall outage probability of the considered system is calculated as

$$P_{out} = 1 - (1 - P_{out}^1)(1 - P_{out}^2) \quad (31)$$

where P_{out}^1 and P_{out}^2 are provided in (15) and (30), respectively.

IV. ERGODIC CAPACITY ANALYSIS

In this section, the ergodic capacity analysis is presented, first for the near user U_1 and then for the far user U_2 .

A. Ergodic Rate from BS to U_1

For the near user U_1 , the ergodic capacity is given as [49]

$$\begin{aligned} R_{ave}^1 &= E \left[\frac{1}{2} \log_2 (1 + \gamma_1) \right] \\ &= \int_0^\infty \frac{1}{2} \log_2 (1 + \gamma) f_{\gamma_1}(\gamma) d\gamma \\ &= \frac{1}{2 \ln 2} \int_0^\infty \frac{1 - F_{\gamma_1}(\gamma)}{1 + \gamma} d\gamma. \end{aligned} \quad (32)$$

By using (5), $F_{\gamma_1}(\gamma)$ is calculated as

$$\begin{aligned} F_{\gamma_1}(\gamma) &= \mathcal{P} \left\{ \frac{G^2 |h_1|^2 |h_r|^2 P_1}{G^2 |h_1|^2 N_0 + N_0} < \gamma \right\} \\ &= \mathcal{P} \{ |h_1|^2 G^2 (|h_r|^2 P_1 - N_0 \gamma) < N_0 \gamma \}. \end{aligned} \quad (33)$$

From (33), it is observed that $F_{\gamma_1}(\gamma)$ may be written as

$$\begin{aligned} F_{\gamma_1}(\gamma) &= \mathcal{P} \left\{ |h_r|^2 > \frac{N_0 \gamma}{P_1}, |h_1|^2 < \frac{N_0 \gamma / G^2}{|h_r|^2 P_1 - N_0 \gamma} \right\} \\ &+ \mathcal{P} \left\{ |h_r|^2 < \frac{N_0 \gamma}{P_1} \right\} \end{aligned} \quad (34)$$

which is expressed in integral form as

$$\begin{aligned} F_{\gamma_1}(\gamma) &= \int_{\frac{N_0 \gamma}{P_1}}^\infty f_{|h_r|^2}(h) F_{|h_1|^2} \left(\frac{N_0 \gamma / G^2}{h P_1 - N_0 \gamma} \right) dh \\ &+ \int_0^{\frac{N_0 \gamma}{P_1}} f_{|h_r|^2}(h) dh. \end{aligned} \quad (35)$$

Then, by substituting (13) in (35) and after some arrangements $F_{\gamma_1}(\gamma)$ is written as,

$$F_{\gamma_1}(\gamma) = 1 - J_1 + J_2 \quad (36)$$

where

$$J_1 = \int_{\frac{N_0 \gamma}{P_1}}^\infty f_{|h_r|^2}(h) \omega_1 \left(\frac{N_0 \gamma / G^2}{h P_1 - N_0 \gamma}, \bar{\gamma}_1 \right) dh \quad (37)$$

while

$$J_2 = \int_{\frac{N_0 \gamma}{P_1}}^\infty f_{|h_r|^2}(h) \omega_2 \left(\frac{N_0 \gamma / G^2}{h P_1 - N_0 \gamma}, \bar{\gamma}_1 \right) dh. \quad (38)$$

Then, by using (10) and (14a), (37) becomes

$$\begin{aligned} J_1 &= \int_{\frac{N_0 \gamma}{P_1}}^\infty \frac{1}{\bar{\gamma}_r} e^{-\frac{h}{\bar{\gamma}_r}} 4 \sqrt{\frac{1}{\bar{\gamma}_1} \frac{N_0 \gamma / G^2}{h P_1 - N_0 \gamma}} \\ &\times K_1 \left(4 \sqrt{\frac{1}{\bar{\gamma}_1} \frac{N_0 \gamma / G^2}{h P_1 - N_0 \gamma}} \right) dh. \end{aligned} \quad (39)$$

Applying $t = h P_1 - N_0 \gamma$ transformation, (39) is written as

$$\begin{aligned} J_1 &= \kappa_2 \sqrt{\frac{N_0 \gamma / G^2}{\bar{\gamma}_1}} \\ &\times \int_0^\infty t^{-1/2} e^{-\frac{t}{\bar{\gamma}_r P_1}} K_1 \left(2 \sqrt{\frac{1}{\bar{\gamma}_1} \frac{N_0 \gamma / G^2}{t}} \right) dt \end{aligned} \quad (40)$$

where $\kappa_2 = \frac{4}{\bar{\gamma}_r P_1} e^{-\frac{N_0 \gamma}{\bar{\gamma}_r P_1}}$. After that, by using [47, 11] and [45, 9.31-1,2,4,5], [47, 14] and [45, 9.31-2] and [45, 9.31-5], (40) is solved as

$$J_1 = 2e^{-\frac{N_0 \gamma}{\bar{\gamma}_r P_1}} G_{2,5}^{3,2} \left(\frac{N_0 \gamma / G^2}{\bar{\gamma}_r \bar{\gamma}_1 P_1} \middle| \begin{matrix} 2, \frac{3}{2} \\ 1, 1, 0, \frac{3}{2}, 2 \end{matrix} \right). \quad (41)$$

Furthermore, by substituting (10) and (14b) in (38) and using $t = h P_1 - N_0 \gamma$ transformation, J_2 is expressed as

$$\begin{aligned} J_2 &= \kappa_2 \frac{N_0 \gamma / G^2}{\bar{\gamma}_1} \\ &\times \int_0^\infty t^{-1} e^{-\frac{t}{\bar{\gamma}_r P_1}} \left[K_1 \left(2 \sqrt{\frac{1}{\bar{\gamma}_1} \frac{N_0 \gamma / G^2}{t}} \right) \right]^2 dt. \end{aligned} \quad (42)$$

Then, applying [48, 03.04.26.0014.01], [45, 9.31-5], [45, 9.31-1], [45, 9.31-2] and [45, 7.813-1] respectively, J_2 is derived as

$$J_2 = \frac{\sqrt{\pi}}{2} e^{-\frac{N_0 \gamma}{\bar{\gamma}_r P_1}} G_{5,2}^{1,4} \left(\frac{\bar{\gamma}_1 \bar{\gamma}_r P_1}{4 N_0 \gamma / G^2} \middle| \begin{matrix} 0,0,1,-1,0 \\ 0,-\frac{1}{2} \end{matrix} \right). \quad (43)$$

Besides, by using (41) and (43) in (32), the ergodic capacity for the near user U_1 is calculated as

$$R_{ave}^1 = \frac{1}{2 \ln 2} \int_0^\infty \frac{J_1 - J_2}{1 + \gamma} d\gamma = \frac{1}{2 \ln 2} [\mathcal{M}_1 - \mathcal{M}_2] \quad (44)$$

where

$$\mathcal{M}_1 = 2 \int_0^\infty \frac{e^{-\frac{N_0 \gamma}{\bar{\gamma}_r P_1}}}{1 + \gamma} G_{2,5}^{3,2} \left(\frac{N_0 \gamma / G^2}{\bar{\gamma}_r \bar{\gamma}_1 P_1} \middle| \begin{matrix} 2, \frac{3}{2} \\ 1, 1, 0, \frac{3}{2}, 2 \end{matrix} \right) d\gamma, \quad (45)$$

and

$$\mathcal{M}_2 = \frac{\sqrt{\pi}}{2} \int_0^\infty \frac{e^{-\frac{N_0 \gamma}{\bar{\gamma}_r P_1}}}{1 + \gamma} G_{5,2}^{1,4} \left(\frac{\bar{\gamma}_1 \bar{\gamma}_r P_1}{4 N_0 \gamma / G^2} \middle| \begin{matrix} 0,0,1,-1,0 \\ 0,-\frac{1}{2} \end{matrix} \right) d\gamma. \quad (46)$$

Using Meijer's G representations of polynomial and exponential functions as $(1+x)^\alpha = \frac{1}{\Gamma(-\alpha)} G_{1,1}^{1,1} \left(x \middle| \begin{smallmatrix} 1+\alpha \\ 0 \end{smallmatrix} \right)$ [47, 10] and $e^{-x} = G_{0,1}^{1,0} \left(x \middle| \begin{smallmatrix} - \\ 0 \end{smallmatrix} \right)$ [47, 11], (45) can be expressed as

$$\mathcal{M}_1 = 2 \int_0^\infty G_{1,1}^{1,1} \left(\gamma \middle| \begin{smallmatrix} 0 \\ 0 \end{smallmatrix} \right) G_{0,1}^{1,0} \left(\frac{N_0 \gamma}{\bar{\gamma}_r P_1} \middle| \begin{smallmatrix} - \\ 0 \end{smallmatrix} \right) \times G_{2,5}^{3,2} \left(\frac{N_0 \gamma / G^2}{\bar{\gamma}_r \bar{\gamma}_1 P_1} \middle| \begin{smallmatrix} 2, \frac{3}{2} \\ 1, 1, 0, \frac{3}{2}, 2 \end{smallmatrix} \right) d\gamma \quad (47)$$

and with the help of [48, 07.34.21.0081.01], \mathcal{M}_1 is obtained as

$$\mathcal{M}_1 = 2G_{1,1:0:3,2}^{1,1:1:0:3,2} \left(\frac{N_0}{\bar{\gamma}_r P_1}, \frac{N_0 / G^2}{\bar{\gamma}_r \bar{\gamma}_1 P_1} \middle| \begin{smallmatrix} 0 \\ 0 \end{smallmatrix} \middle| \begin{smallmatrix} 2, \frac{3}{2} \\ 1, 1, 0, \frac{3}{2}, 2 \end{smallmatrix} \right) \quad (48)$$

in closed-form where $G_{p,q;p_1,q_1:p_2,q_2}^{m,n;m_1,n_1;m_2,n_2} (\cdot, \cdot \middle| \cdot \middle| \cdot)$ is the extended generalized bi-variate Meijer's G function [50]. It should be noted that the well-known numeric computing environments have not yet included the bi-variate Meijer's G as a built-in function, but several algorithms for its calculation are available in the literature [51].

Further, by following similar steps, \mathcal{M}_2 is written as

$$\mathcal{M}_2 = \frac{\sqrt{\pi}}{2} \int_0^\infty G_{1,1}^{1,1} \left(\gamma \middle| \begin{smallmatrix} 0 \\ 0 \end{smallmatrix} \right) G_{0,1}^{1,0} \left(\frac{N_0 \gamma}{\bar{\gamma}_r P_1} \middle| \begin{smallmatrix} - \\ 0 \end{smallmatrix} \right) \times G_{5,2}^{1,4} \left(\frac{\bar{\gamma}_1 \bar{\gamma}_r P_1}{4N_0 \gamma / G^2} \middle| \begin{smallmatrix} 0, 0, 1, -1, 0 \\ 0, -\frac{1}{2} \end{smallmatrix} \right) d\gamma \quad (49)$$

and using [48, 07.34.21.0081.01], it is derived as

$$\mathcal{M}_2 = \frac{\sqrt{\pi}}{2} G_{1,1:0:4,1}^{1,1:1:0:4,1} \left(\frac{N_0}{\bar{\gamma}_r P_1}, \frac{4N_0 / G^2}{\bar{\gamma}_r \bar{\gamma}_1 P_1} \middle| \begin{smallmatrix} 0 \\ 0 \end{smallmatrix} \middle| \begin{smallmatrix} 1, \frac{3}{2} \\ 1, 1, 0, 2, 1 \end{smallmatrix} \right) \quad (50)$$

in closed-form. Finally, by substituting (48) and (50) in (44), the ergodic capacity expression for the near user U_1 can be obtained in closed-form.

B. Ergodic Rate from BS to U_2

By following steps similar to those in finding out R_{ave}^1 , the ergodic capacity for U_2 is written as

$$R_{ave}^2 = E \left[\frac{1}{2} \log_2 (1 + \gamma_2) \right] = \frac{1}{2 \ln 2} \int_0^\infty \frac{1 - F_{\gamma_2}(\gamma)}{1 + \gamma} d\gamma \quad (51)$$

where $F_{\gamma_2}(\gamma)$ is expressed as follows, with the help of (6),

$$F_{\gamma_2}(\gamma) = \mathcal{P} \left\{ \frac{G^2 |h_2|^2 |h_r|^2 P_2}{G^2 |h_2|^2 |h_r|^2 P_1 + G^2 |h_2|^2 N_0 + N_0} < \gamma \right\} = \mathcal{P} \left\{ |h_r|^2 > \frac{N_0 \gamma}{P_2 - P_1 \gamma}, |h_2|^2 < \frac{N_0 \gamma / G^2}{|h_r|^2 (P_2 - P_1 \gamma) - N_0 \gamma} \right\} + \mathcal{P} \left\{ |h_r|^2 < \frac{N_0 \gamma}{P_2 - P_1 \gamma} \right\} \quad (52)$$

and from the requirement $\gamma < P_2 / P_1$, (52) becomes

$$F_{\gamma_2}(\gamma) = \int_0^{\frac{N_0 \gamma}{\beta}} f_{|h_r|^2}(h) dh + \int_{\frac{N_0 \gamma}{\beta}}^\infty f_{|h_r|^2}(h) F_{|h_2|^2} \left(\frac{N_0 \gamma / G^2}{h \beta - N_0 \gamma} \right) dh \quad (53)$$

where $\beta = P_2 - P_1 \gamma$. Then, by using the CDF of the ordered channel gain for U_2 , given by (27), (53) takes the form

$$F_{\gamma_2}(\gamma) = 1 - J_3 \quad (54)$$

where J_3 is

$$J_3 = \int_{\frac{N_0 \gamma}{\beta}}^\infty f_{|h_r|^2}(h) \omega_2 \left(\frac{N_0 \gamma / G^2}{\beta h - N_0 \gamma}, \bar{\gamma}_2 \right) dh. \quad (55)$$

Thus, R_{ave}^2 is written as

$$R_{ave}^2 = \frac{1}{2 \ln 2} \int_0^{\frac{P_2}{P_1}} \frac{J_3}{1 + \gamma} d\gamma. \quad (56)$$

It is worth pointing out that J_3 has an integral expression similar to J_2 with the replacement of the parameters $\bar{\gamma}_1$ with $\bar{\gamma}_2$, and P_1 with β in (38). Therefore, by following steps similar to those in deriving J_2 , J_3 is obtained as

$$J_3 = J_2|_{\bar{\gamma}_1 \rightarrow \bar{\gamma}_2, P_1 \rightarrow \beta} = \frac{\sqrt{\pi}}{2} e^{-\frac{N_0 \gamma}{\bar{\gamma}_r \beta}} G_{5,2}^{1,4} \left(\frac{\bar{\gamma}_2 \bar{\gamma}_r \beta}{4N_0 \gamma / G^2} \middle| \begin{smallmatrix} 0, 0, 1, -1, 0 \\ 0, -\frac{1}{2} \end{smallmatrix} \right). \quad (57)$$

Finally, by substituting (57) in (56), the ergodic capacity at the far user is obtained as

$$R_{ave}^2 = \frac{1}{2 \ln 2} \frac{\sqrt{\pi}}{2} \int_0^{\frac{P_2}{P_1}} \frac{1}{1 + \gamma} e^{-\frac{N_0 \gamma}{\bar{\gamma}_r \beta}} \times G_{5,2}^{1,4} \left(\frac{\bar{\gamma}_2 \bar{\gamma}_r \beta}{4N_0 \gamma / G^2} \middle| \begin{smallmatrix} 0, 0, 1, -1, 0 \\ 0, -\frac{1}{2} \end{smallmatrix} \right) d\gamma. \quad (58)$$

Please note that the single integral form in (58) is not analytically tractable. Therefore, it should be evaluated numerically.

C. Overall Ergodic Rate

For the proposed system, the overall achievable ergodic capacity is calculated as

$$R_{ave} = R_{ave}^1 + R_{ave}^2. \quad (59)$$

where R_{ave}^1 and R_{ave}^2 are provided in (44) and (58), respectively.

V. NUMERICAL RESULTS AND SIMULATIONS

In this section, the performance of the proposed NOMA-enabled cooperative V2V system is presented through numerical calculations of the analytic derivations by comparing simulations for miscellaneous scenarios. During the calculations and Monte-Carlo trials, it is assumed that all nodes, BS, R, U_1 , and U_2 , align on a straight line and the noise variance at each node is set to $N_0 = 1$. Without loss of generality, the average channel gains are considered to be associated with the corresponding distances between the nodes as $\Omega_{sr} = 1/(1 + d_{sr}^2)$, $\Omega_1 = 1/(1 + d_1^2)$ and $\Omega_2 = 1/(1 + d_2^2)$. Additionally, the obtained results of the proposed system are compared with the simulation results of the traditional cooperative OMA (COMA) technique [52] to provide better and comprehensive insight into the performance of the NOMA technique. Herein, the received SNRs of using COMA technique considering (3) are calculated as $\gamma_i^{COMA} =$

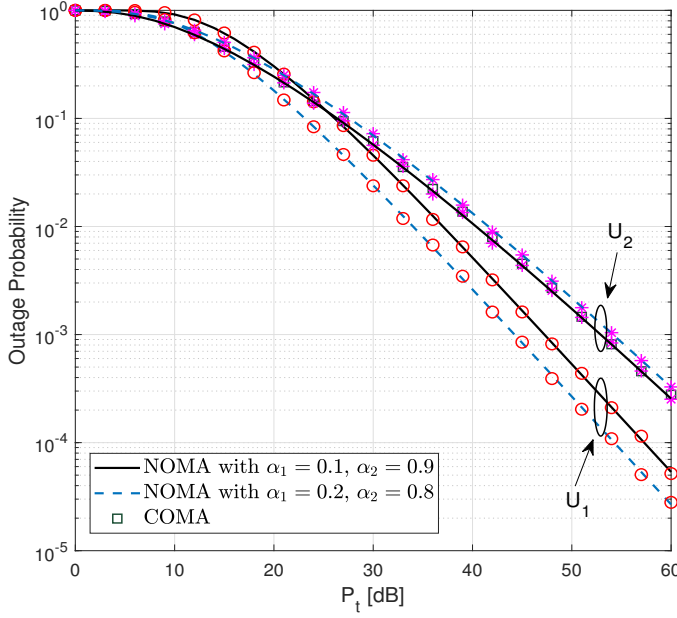


Fig. 2. The outage performance of the near and far users in the proposed NOMA-enabled vehicular communication system and comparison with the COMA. $\mathcal{R}_1 = \mathcal{R}_2 = 0.5$ bit/Hz/s and $(d_{sr}, d_1, d_2) = (0.5, 0.5, 0.5)$.

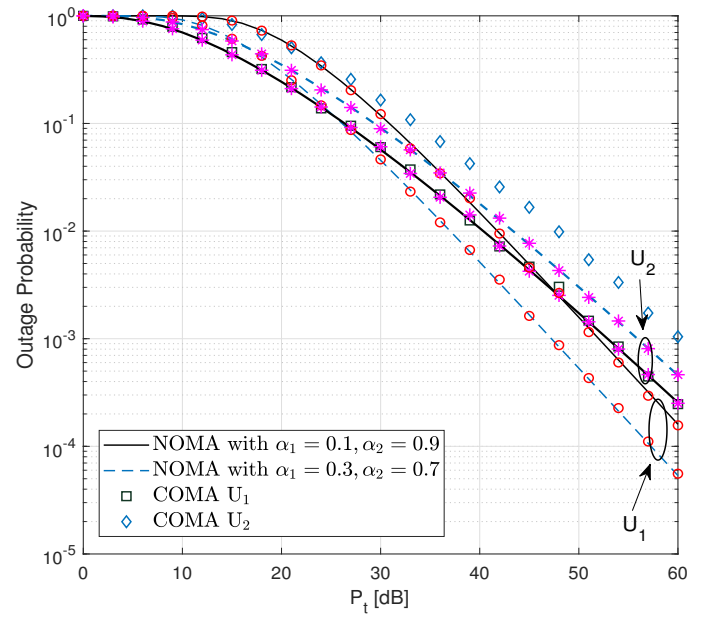


Fig. 3. The outage performance comparison of NOMA and COMA techniques over double Rayleigh fading channels. $\mathcal{R}_1 = 1$ bit/Hz/s, $\mathcal{R}_2 = 0.5$ bit/Hz/s and $(d_{sr}, d_1, d_2) = (0.5, 0.5, 0.5)$.

$G|h_i|^2|h_r|^2P_t/(G|h_i|^2N_0 + N_0)$ and the corresponding rate is given as $\mathcal{R}_i^{COMA} = 0.25 \log_2(1 + \gamma_i^{COMA})$ where $i \in \{1, 2\}$ for the near and far users, respectively. Please note that, the rate of applying the COMA technique is half the rate of using the NOMA technique since the symbol of users are not transmitted simultaneously in the COMA scheme [2]. Moreover, for a fair comparison between COMA and NOMA, we assume that each user is allocated equal bandwidth and equal transmission power. Furthermore, the presented figures show the analytical and Monte-Carlo simulation results together to verify the derivations, where the markers on the plots represent simulation outcomes while the lines show numerical results of the analytic expressions.

Fig. 2 shows the variation of the outage probabilities for both users with respect to the total transmitting power P_t for two different power allocation regimes at the BS, $\alpha_1 = 0.1$ ($\alpha_2 = 0.9$) and $\alpha_1 = 0.2$ ($\alpha_2 = 0.8$). Here, the simulation results of the COMA scheme are also provided for the sake of comparison. In the proposed system, since the symbol transmission from BS to the users takes two time slots and by assuming that BS uses the same modulation scheme for both users, let us say a binary modulation scheme, target data rates are assigned as $\mathcal{R}_1 = \mathcal{R}_2 = 0.5$ bit/Hz/s. Further, the normalized distances are chosen as $d_{sr} = 0.5$, $d_1 = 0.5$ and $d_2 = 0.5$. The figure shows that increasing the power ratio α_1 from 0.1 to 0.2, which means doubling the transmission power allocated to U_1 (i.e., 100% increase), provides 3 dB SNR gain for the outage performance of U_1 with the strong channel. On the other hand, decreasing the power ratio α_2 from 0.9 to 0.8, which means 11.5% reduction at the transmission power allocated to U_2 , results in an approximately 1.5 dB SNR loss at the outage probability performance of U_2 with

the weak channel. This tells us that the effective gain of U_1 becomes greater than the effective loss of U_2 as the P_2/P_1 ratio decreases and therefore overall system performance can also be enhanced in this way. Furthermore, Fig. 2 presents the performance of the COMA scheme and it reveals that the outage performance of the NOMA scheme at U_1 outperforms the COMA's since the channel gain of U_1 is greater than that of U_2 . On the other hand, it is shown that the performance of COMA for U_2 becomes slightly better with the increasing α_1 value. However, it is worth noting that the NOMA scheme will offer better spectral efficiency and user fairness than COMA as multiple users are operated simultaneously.

The outage probability performances of U_1 and U_2 with respect to the total transmitting power P_t are depicted in comparison with the COMA scheme in Fig. 3. Here the target data rates are chosen as $\mathcal{R}_1 = 1$ bit/Hz/s and $\mathcal{R}_2 = 0.5$ bit/Hz/s while the distances are assumed to be $d_{sr} = 0.5$, $d_1 = 0.5$ and $d_2 = 0.5$. Comparing the performance plot of U_1 in Fig. 2 for $\alpha_1 = 0.1$ and $\mathcal{R}_1 = 0.5$ bit/Hz/s with that in Fig. 3 for $\alpha_1 = 0.3$ and $\mathcal{R}_1 = 1$ bit/Hz/s, it is observed that U_1 has similar outage performance, for example 9×10^{-3} and 6×10^{-3} at 50 dB. Therefore, by allocating more power, U_1 can double its target data rate with a negligible loss in the outage performance. On the other hand, as expected, the outage performance of U_2 gets worse with increasing α_1 . Furthermore, it is seen from the figure that the performance of U_2 surpasses U_1 for $\alpha_1 = 0.1$ at low and middle SNR regions due to the substantially low transmit power of the BS for U_1 when $P_t < 45$ dB. The figure also shows that the outage performance of U_1 with COMA technique is better than the NOMA-enabled transmitting scheme when $\alpha_1 = 0.1$ for SNR values up to 45 dB. This is due to the fact that since the

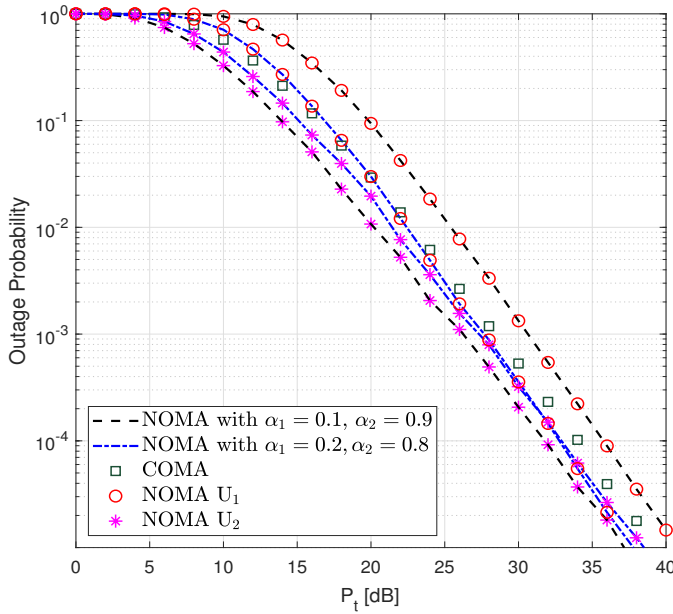


Fig. 4. The outage performance of the near and far users for using Nakagami- m fading for all channels in NOMA-enabled vehicular communication system. $(d_{sr}, d_1, d_2) = (0.5, 0.5, 0.5)$, $m = 2$ and $\mathcal{R}_1 = \mathcal{R}_2 = 0.5$ bit/Hz/s

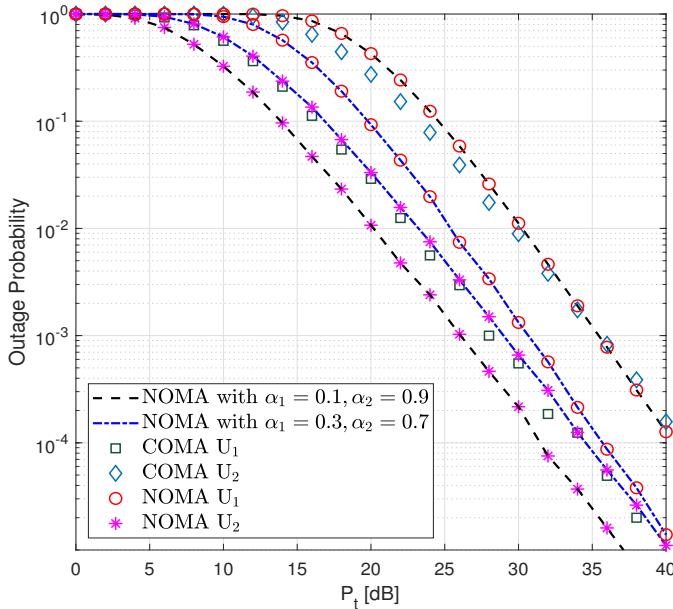


Fig. 5. The outage performance of the near and far users for using Nakagami- m fading for all channels in NOMA-enabled vehicular communication system. $(d_{sr}, d_1, d_2) = (0.5, 0.5, 0.5)$, $m = 2$, $\mathcal{R}_1 = 1$ bit/Hz/s and $\mathcal{R}_2 = 0.5$ bit/Hz/s

rate of U_1 in COMA is half of the NOMA and more power is allocated to U_1 in COMA, outage probability outperforms compared to NOMA at low SNR. Besides, NOMA shows better performance for U_1 at high SNR when $\alpha_1 = 0.1$ since its performance is exponentially increased by the increasing transmit power of the BS considering (5). However, it is obvious that the NOMA technique will provide better spectral

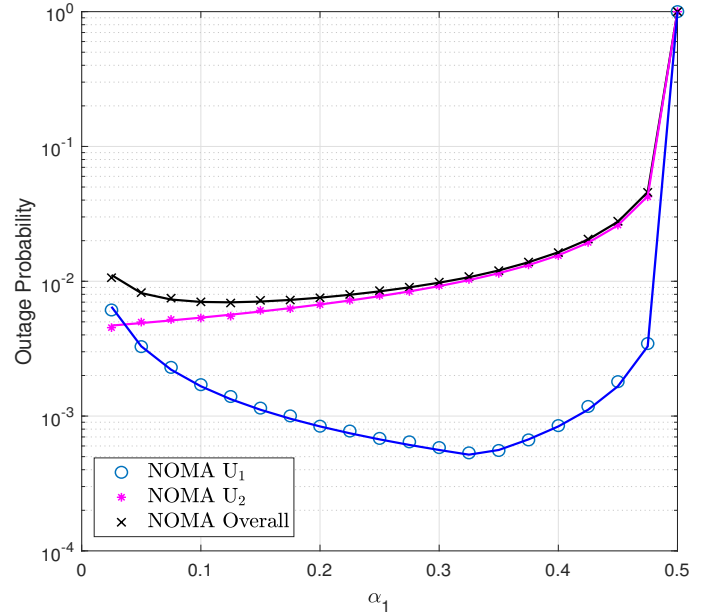


Fig. 6. The outage probability performances of the overall system and the users in the NOMA-enabled cooperative vehicular communication system. $\mathcal{R}_1 = \mathcal{R}_2 = 0.5$ bit/Hz/s, $(d_{sr}, d_1, d_2) = (0.5, 0.5, 0.8)$ and $P_t = 45$ dB.

efficiency than the COMA method since two users are assisted simultaneously in the proposed system.

Fig. 4 presents the performance of the proposed system using the same system parameters in Fig. 2 over Nakagami- m fading channels to provide a comparison with the double Rayleigh fading channels. For a fair comparison, we set the Nakagami- m parameter of both channels to the same value as $m = 2$. Since the fading conditions in Nakagami- m channels are better than those in the double Rayleigh channels, power allocation among the users affects their performances that yields the performance of U_2 becomes better than U_1 's performance. It is seen from Fig. 2 that, an increase in α_1 from 0.1 to 0.2 results in an enhancement in the U_1 's system performance while the system performance of U_2 degrades. Moreover, when considering the Nakagami- m fading channels, system performance of both users act as similar as in the case of the double Rayleigh fading channels. Furthermore, it is seen from Fig. 4 that the approximately 26 dB SNR gain is achieved for the outage probability of 10^{-3} when assuming the Nakagami- m fading channels for COMA.

Fig. 5 shows the variation of the outage probability with P_t in the proposed system where the channels are subject to Nakagami- m fading, instead of the double Rayleigh fading. Here again, the Nakagami- m parameter values for both channels are assigned the same as $m = 2$. Apart from Fig. 3, it is seen from Fig. 5 that U_2 outperforms U_1 in terms of system performance, since more power is allocated for U_2 . Contrary to the observed results in Fig. 3, the increasing α_1 causes the increasing outage probability for U_1 while U_2 's outage probability decreases. Furthermore, when employing COMA scheme over Nakagami- m fading channels, approximately 25 dB SNR gain for U_1 and U_2 is achieved to have the outage

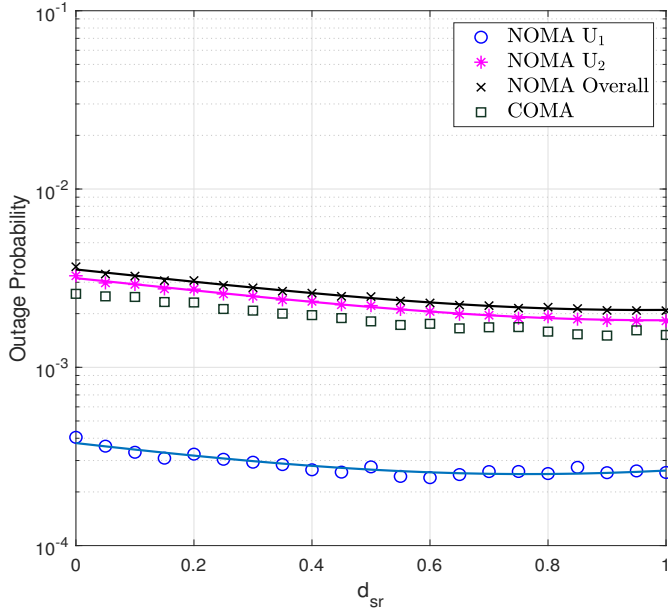


Fig. 7. The outage performance comparison of NOMA and COMA techniques with respect to the distance between BS and the relay. $(\alpha_1, \alpha_2) = (0.2, 0.8)$, $\mathcal{R}_1 = 0.5$ bit/Hz/s, $\mathcal{R}_2 = 0.5$ bit/Hz/s, and $P_t = 50$ dB.

probability of 10^{-3} when comparing with the double Rayleigh fading channels.

The outage probability performances of both users and the overall system performance with respect to the power allocation parameter α_1 is presented in Fig. 6. Here, $P_t = 45$ dB and the normalized distances are $d_{sr} = d_1 = 0.5$ and $d_2 = 0.8$. The figure shows that the outage probability performance of the overall system and U_2 converge to each other in the range $0.32 \leq \alpha_1 < 0.5$. It is also seen that $\alpha_1 = 0.32$ value, which is the starting point for the convergence of the outage probabilities for U_2 and the overall system, is the optimum power allocation coefficient value for U_1 . The figure also reveals that increasing the allocated power for U_1 causes degradation in the power level of U_2 , since U_1 's performance also degrades due to the SIC procedure failure. Further, it can be observed that for $\alpha_1 > 0.5$, the criteria $\gamma_{th2} < P_2/P_1$ in (8) cannot be satisfied and the outage probabilities rapidly reach 1.

Fig. 7 depicts the outage probabilities of U_1 , U_2 and the overall system versus distance d_{sr} . Here, the system parameters are set to $P_t = 50$ dB, $\alpha_1 = 0.2$ ($\alpha_2 = 0.8$), $\mathcal{R}_1 = 0.5$ bit/Hz/s, and $\mathcal{R}_2 = 0.5$ bit/Hz/s. Furthermore, the distances between BS- U_1 and BS- U_2 are assigned as $d = 1$. However, R is in motion under the assumption $d_1 = d_2 = d - d_{sr}$. From the figure, it is observed that U_1 has a better performance compared to U_2 , since the channel gain of U_1 is greater than that of U_2 . Therefore, NOMA outperforms COMA for the strong user U_1 . On the other hand, as an important result, it is seen that COMA has slightly better outage performance than NOMA for the weak user U_2 . However, NOMA still has the advantages of providing better spectral efficiency and user fairness as multiple users use the same frequency band at the

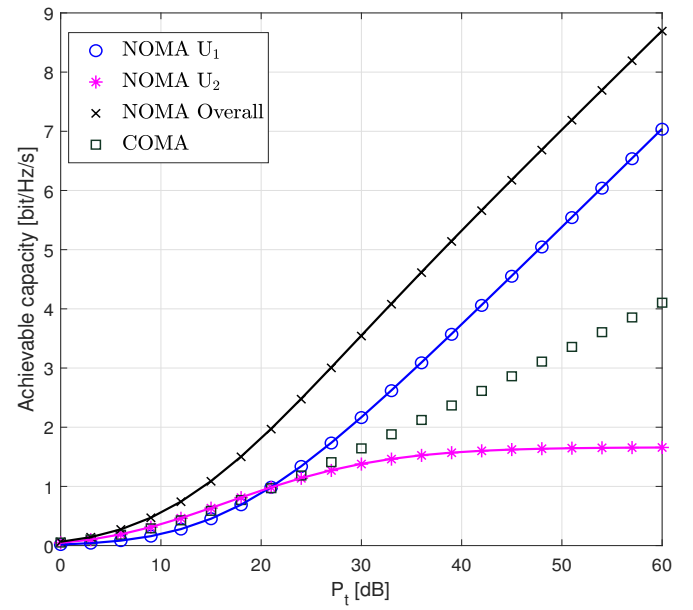


Fig. 8. The comparison of ergodic capacity of the users and the overall system for COMA and NOMA for $(d_{sr}, d_1, d_2) = (0.5, 0.5, 0.5)$ where for NOMA $(\alpha_1, \alpha_2) = (0.1, 0.9)$.

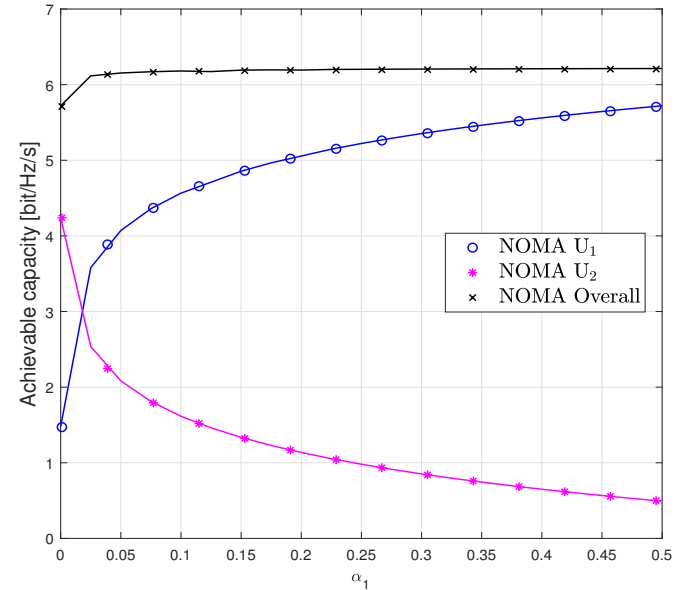


Fig. 9. The ergodic capacities of the near and far users and of the overall system. $P_t = 45$ dB and $(d_{sr}, d_1, d_2) = (0.5, 0.5, 0.8)$.

same time.

For the proposed system, achievable ergodic capacity for the overall system and also the ergodic capacities for U_1 and U_2 are presented in Fig. 8 for NOMA and COMA schemes, with respect to P_t for the case of $\alpha_1 = 0.1$ ($\alpha_2 = 0.9$). The figure reveals that increasing P_t results in saturation at the achievable rate performance of U_2 at high SNR due to interference from U_1 . However, since U_1 is able to perfectly

decode its information without an interference from U_2 after a successful SIC operation, its own and therefore the overall system performance increase. It is also shown that U_1 reaches higher achievable capacity values when using NOMA technique compared to the case of employing traditional COMA scheme while the achievable capacity for U_2 for COMA technique is better than that for the NOMA method.

Finally, the ergodic capacity performances of the proposed system is given in Fig. 9 for changing α_1 values when $P_t = 45$ dB and the normalized distances are assigned as $d_{sr} = 0.5$, $d_1 = 0.5$ and $d_2 = 0.8$. The figure shows that the overall achievable data rate saturates to approximately 6 bit/Hz/s even for small values of α_1 while increasing α_1 provides higher achievable data rates for U_1 . However, the achievable data rate for U_2 decreases in this case.

VI. CONCLUSION

In this paper, performance of the NOMA-enabled cooperative vehicular communication systems, where the transmission from the base station to the end-users has been established with the help of a relay employing fixed-gain AF relaying scheme has been investigated. Since the fading channels in vehicular communication systems exhibit cascaded fading characteristics, theoretical analyses have been carried out by considering the double Rayleigh fading channel model as a realistic fading scheme. Moreover, the results are compared with the Nakagami- m fading. During the analyses, we first obtained closed-form expressions for the outage probabilities of the strong/near and weak/far users; after that, the exact ergodic capacities for these users have been derived under double Rayleigh fading. Then we provided the numerical results of the calculations of analytic expressions for different system parameters and verified them by Monte-Carlo simulations. We also provided the simulation results for the COMA scheme and compared the performances of NOMA and COMA methods in the proposed V2V communication system. The results show that the near user has better performance in the NOMA-enabled systems compared to the systems employing the COMA technique due to the SIC being perfectly applied on the near user. On the other hand, the performance of the weak user worsens in case of using the NOMA technique compared to the COMA method even though the allocated power at the base station for the near user remains at small values.

ACKNOWLEDGMENT

The work of Semiha Koşu was supported by the Research Fund of the Istanbul Technical University under Project MYL-2019-42491.

REFERENCES

- [1] M. Simon and M. Alouini, *Digital Communication over Fading Channels*. Wiley, 2005.
- [2] Y. Saito, Y. Kishiyama, A. Benjebbour, T. Nakamura, A. Li, and K. Higuchi, "Non-orthogonal Multiple Access (NOMA) for Cellular Future Radio Access," in *Proc. IEEE Veh. Tech. Conf.*, 2013, pp. 1–5.
- [3] S. A. A. Shah, E. Ahmed, M. Imran, and S. Zeadally, "5G for Vehicular Communications," *IEEE Commun. Magazine*, vol. 56, no. 1, pp. 111–117, 2018.
- [4] Y. Yang and K. Hua, "Emerging Technologies for 5G-Enabled Vehicular Networks," *IEEE Access*, vol. 7, pp. 181–117, 2019.
- [5] G. Luo, Q. Yuan, H. Zhou, N. Cheng, Z. Liu, F. Yang, and X. S. Shen, "Cooperative Vehicular Content Distribution in Edge Computing Assisted 5G-VANET," *China Commun.*, vol. 15, no. 7, pp. 1–17, 2018.
- [6] J. Lee and B. Park, "Development and Evaluation of a Cooperative Vehicle Intersection Control Algorithm Under the Connected Vehicles Environment," *IEEE Trans. on Intelligent Transportation Systems*, vol. 13, no. 1, pp. 81–90, 2012.
- [7] S. K. Gehrig and F. J. Stein, "Collision Avoidance for Vehicle-Following Systems," *IEEE Trans. on Intelligent Transportation Systems*, vol. 8, no. 2, pp. 233–244, 2007.
- [8] S. Darbha, S. Konduri, and P. R. Pagilla, "Benefits of V2V Communication for Autonomous and Connected Vehicles," *IEEE Trans. on Intelligent Transportation Systems*, vol. 20, no. 5, pp. 1954–1963, 2019.
- [9] Y. Ibdah and Y. Ding, "Statistical properties for Cascaded Rayleigh Fading Channel Models," in *Proc. 9th Int. Conf. on Inf., Commun. Signal Process.*, 2013, pp. 1–5.
- [10] V. Erceg, S. J. Fortune, J. Ling, A. J. Rustako, and R. A. Valenzuela, "Comparisons of a Computer-based Propagation Prediction Tool with Experimental Data Collected in Urban Microcellular Environments," *IEEE Journal on Selected Areas in Commun.*, vol. 15, no. 4, pp. 677–684, 1997.
- [11] D. W. Matolak and J. Frolik, "Worse-than-Rayleigh fading: Experimental Results and Theoretical Models," *IEEE Commun. Magazine*, vol. 49, no. 4, pp. 140–146, 2011.
- [12] V. Erceg, S. Fortune, J. Ling, A. Rustako, and R. Valenzuela, "Comparisons of a computer-based propagation prediction tool with experimental data collected in urban microcellular environments," *IEEE Journal on Selected Areas in Communications*, vol. 15, no. 4, pp. 677–684, 1997.
- [13] M. Seyfi, S. Muhaidat, J. Liang, and M. Uysal, "Relay Selection in Dual-Hop Vehicular Networks," *IEEE Signal Process. Lett.*, vol. 18, no. 2, pp. 134–137, Feb. 2011.
- [14] D. Qin, Y. Wang, and T. Zhou, "Performance Analysis of Bidirectional AF Based Cooperative Vehicular Networks," *IEEE Trans. on Veh. Tech.*, vol. 69, no. 2, pp. 2274–2279, Feb. 2020.
- [15] Y. Alghorani, G. Kaddoum, S. Muhaidat, and S. Pierre, "On the Approximate Analysis of Energy Detection Over n -Rayleigh Fading Channels Through Cooperative Spectrum Sensing," *IEEE Wireless Commun. Lett.*, vol. 4, no. 4, pp. 413–416, Aug. 2015.
- [16] A. Al Hammadi, O. Alhussein, P. C. Sofotasios, S. Muhaidat, M. Al-Qutayri, S. Al-Araji, G. K. Karagiannidis, and J. Liang, "Unified Analysis of Cooperative Spectrum Sensing Over Composite and Generalized Fading Channels," *IEEE Trans. on Veh. Tech.*, vol. 65, no. 9, pp. 6949–6961, Sep. 2016.
- [17] P. S. Bithas, G. P. Efthymoglou, and A. G. Kanatas, "V2V Cooperative Relaying Communications Under Interference and Outdated CSI," *IEEE Trans. on Veh. Tech.*, vol. 67, no. 4, pp. 3466–3480, April 2018.
- [18] S. Wang, D. Wang, C. Li, and W. Xu, "Full Duplex AF and DF Relaying Under Channel Estimation Errors for V2V Communications," *IEEE Access*, vol. 6, pp. 65 321–65 332, 2018.
- [19] O. Abbasi, A. Ebrahimi, and N. Mokari, "NOMA Inspired Cooperative Relaying System Using an AF Relay," *IEEE Wireless Commun. Lett.*, vol. 8, no. 1, pp. 261–264, Feb. 2019.
- [20] M. Tian, S. Zhao, Q. Li, and J. Qin, "Secrecy Sum Rate Optimization in Nonorthogonal Multiple Access AF Relay Networks," *IEEE Systems Journal*, vol. 13, no. 3, pp. 2712–2715, Sep. 2019.
- [21] Y. Li, Y. Li, Y. Chen, Y. Ye, and H. Zhang, "Performance Analysis of Cooperative NOMA with a Shared AF Relay," *IET Commun.*, vol. 12, no. 19, pp. 2438–2447, 2018.
- [22] Z. Wang, X. Yue, and Z. Peng, "Full-Duplex User Relaying for NOMA System With Self-Energy Recycling," *IEEE Access*, vol. 6, pp. 67 057–67 069, 2018.
- [23] D. Tran, D. Ha, V. N. Vo, C. So-In, H. Tran, T. G. Nguyen, Z. A. Baig, and S. Sanguanpong, "Performance Analysis of DF/AF Cooperative MISO Wireless Sensor Networks With NOMA and SWIPT Over Nakagami- m Fading," *IEEE Access*, vol. 6, pp. 56 142–56 161, 2018.
- [24] Q. Y. Liau, C. Y. Leow, and Z. Ding, "Amplify-and-Forward Virtual Full-Duplex Relaying-Based Cooperative NOMA," *IEEE Wireless Commun. Lett.*, vol. 7, no. 3, pp. 464–467, June 2018.
- [25] D. Wan, M. Wen, F. Ji, Y. Liu, and Y. Huang, "Cooperative NOMA Systems With Partial Channel State Information Over Nakagami- m Fading Channels," *IEEE Trans. on Commun.*, vol. 66, no. 3, pp. 947–958, March 2018.
- [26] Z. Yang, Z. Ding, Y. Wu, and P. Fan, "Novel Relay Selection Strategies for Cooperative NOMA," *IEEE Trans. on Veh. Tech.*, vol. 66, no. 11, pp. 10 114–10 123, Nov. 2017.

- [27] X. Liang, Y. Wu, D. W. K. Ng, Y. Zuo, S. Jin, and H. Zhu, "Outage Performance for Cooperative NOMA Transmission with an AF Relay," *IEEE Commun. Lett.*, vol. 21, no. 11, pp. 2428–2431, Nov. 2017.
- [28] X. Yue, Y. Liu, S. Kang, and A. Nallanathan, "Performance Analysis of NOMA With Fixed Gain Relaying Over Nakagami- m Fading Channels," *IEEE Access*, vol. 5, pp. 5445–5454, 2017.
- [29] Y. Liu, G. Pan, H. Zhang, and M. Song, "Hybrid Decode-Forward Amplify-Forward Relaying With Non-Orthogonal Multiple Access," *IEEE Access*, vol. 4, pp. 4912–4921, 2016.
- [30] R. Jiao, L. Dai, J. Zhang, R. MacKenzie, and M. Hao, "On the Performance of NOMA-Based Cooperative Relaying Systems Over Rician Fading Channels," *IEEE Trans. on Veh. Tech.*, vol. 66, no. 12, pp. 11 409–11 413, Dec. 2017.
- [31] B. E. Y. Belmekki, A. Hamza, and B. Escrig, "On the Outage Probability of Cooperative 5G NOMA at Intersections," in *Proc. IEEE 89th Veh. Tech. Conf. (VTC2019-Spring)*, April 2019, pp. 1–6.
- [32] W. Xie, J. Liao, C. Yu, P. Zhu, and X. Liu, "Physical Layer Security Performance Analysis of the FD-Based NOMA-VC System," *IEEE Access*, vol. 7, pp. 115 568–115 573, 2019.
- [33] B. E. Y. Belmekki, A. Hamza, and B. Escrig, "Outage Analysis of Cooperative NOMA Using Maximum Ratio Combining at Intersections," in *Proc. Int. Conf. on Wireless and Mobile Computing, Net. and Commun. (WiMob)*, Oct. 2019, pp. 1–6.
- [34] T. Kim, Y. Park, H. Kim, and D. Hong, "Cooperative Superposed Transmission in Cellular-Based V2V Systems," *IEEE Trans. on Veh. Tech.*, vol. 68, no. 12, pp. 11 888–11 901, Dec. 2019.
- [35] L. Wei, Y. Chen, D. Zheng, and B. Jiao, "Secure Performance Analysis and Optimization for FD-NOMA Vehicular Communications," *China Commun.*, vol. 17, no. 11, pp. 29–41, Nov. 2020.
- [36] G. Liu, Z. Wang, J. Hu, Z. Ding, and P. Fan, "Cooperative noma broadcasting/multicasting for low-latency and high-reliability 5g cellular v2x communications," *IEEE Internet of Things Journal*, vol. 6, no. 5, pp. 7828–7838, 2019.
- [37] Y. Chen, L. Wang, Y. Ai, B. Jiao, and L. Hanzo, "Performance analysis of noma-sm in vehicle-to-vehicle massive mimo channels," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 12, pp. 2653–2666, 2017.
- [38] N. Jaiswal and N. Purohit, "Performance Evaluation of Non-orthogonal Multiple Access in V2V communications Over Double-Rayleigh Fading Channels," in *Proc. IEEE Conf. on Inf. and Commun. Tech.*, Dec. 2019, pp. 1–5.
- [39] S. Gurugopinath, P. C. Sofotasios, Y. Al-Hammadi, and S. Muhaidat, "Cache-Aided Non-Orthogonal Multiple Access for 5G-Enabled Vehicular Networks," *IEEE Trans. on Veh. Tech.*, vol. 68, no. 9, pp. 8359–8371, Sep. 2019.
- [40] H. Xiao, Y. Chen, S. Ouyang, and A. T. Chronopoulos, "Power Control for Clustering Car-Following V2X Communication System With Non-Orthogonal Multiple Access," *IEEE Access*, vol. 7, pp. 68 160–68 171, 2019.
- [41] A. Pandey and S. Yadav, "Joint impact of nodes mobility and imperfect channel estimates on the secrecy performance of cognitive radio vehicular networks over nakagami-m fading channels," *IEEE Open Journal of Vehicular Technology*, vol. 2, pp. 289–309, 2021.
- [42] Y. Cui, G. Nie, and H. Tian, "Hop progress analysis of two-layer vanets with variant transmission range," *IEEE Wireless Communications Letters*, vol. 8, no. 5, pp. 1473–1476, 2019.
- [43] K. Koufos and C. P. Dettmann, "Outage in motorway multi-lane vanets with hardcore headway distance using synthetic traces," *IEEE Transactions on Mobile Computing*, vol. 20, pp. 2445–2456, 2021.
- [44] J. Men and J. Ge, "Performance Analysis of Non-orthogonal Multiple Access in Downlink Cooperative Network," *IET Commun.*, vol. 9, no. 18, pp. 2267–2273, 2015.
- [45] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*. 7th edition, MA, USA: Elsevier, 2007.
- [46] H. A. David and H. N. Nagaraja, *Order Statistics*. NJ, USA: John Wiley & Sons, Inc., 2003.
- [47] V. S. Adamchik and O. I. Marichev, "The Algorithm for Calculating Integrals of Hypergeometric Type Functions and Its Realization in REDUCE System," in *Proc. Int. Symp. on Symbolic and Algebraic Comput.*, ser. ISSAC '90, 1990, pp. 212–224.
- [48] Wolfram Inc., "The mathematical functions site," <http://functions.wolfram.com>, 2018, [Online].
- [49] J. Men, J. Ge, and C. Zhang, "Performance Analysis of Nonorthogonal Multiple Access for Relaying Networks over Nakagami- m Fading Channels," *IEEE Trans. on Veh. Tech.*, vol. 66, no. 2, pp. 1200–1208, 2017.
- [50] S. C. Gupta, "Integrals Involving Products of G-function," in *Proc. Indian Academy of Sciences - Section A*, 1969, p. 193–200.
- [51] H. Chergui, M. Benjillali, and S. Saoudi, "Performance Analysis of Project-and-Forward Relaying in Mixed MIMO-Pinhole and Rayleigh Dual-Hop Channel," *IEEE Commun. Lett.*, vol. 20, no. 3, pp. 610–613, March 2016.
- [52] Q. Zhang, K. Luo, W. Wang, and T. Jiang, "Joint c-oma and c-noma wireless backhaul scheduling in heterogeneous ultra dense networks," *IEEE Transactions on Wireless Communications*, vol. 19, no. 2, pp. 874–887, Feb 2020.



non-orthogonal multiple access techniques.

Semiha Koşu received the B.S. degree from Anadolu University, Eskişehir, Turkey, in 2018 and the M.Sc. degree in Satellite Communication and Remote Sensing Program from Informatics Institute, Istanbul Technical University, Istanbul, in 2021. She is currently a research assistant and working toward a Ph.D. degree with the Department of Information and Communication Engineering, Istanbul Technical University, Istanbul, Turkey. Her current research interests include the performance analysis of vehicular communications, cooperative communications, and



with Information and Communications Research Group in ITU Informatics Institute. His current research interests are in wireless communications including space-time coding, physical layer network coding, physical layer security, mobile-to-mobile/vehicle-to-vehicle communications and free space optical communications.

Serdar Özgür Ata received the B.Sc., M.Sc., and Ph.D. degrees all in electronics and telecommunication engineering from Istanbul Technical University (ITU), Istanbul, Turkey, in 1996, 2010, and 2017, respectively. He has been chief researcher at the Scientific and Technological Research Council of Turkey - Informatics and Information Security Research Center (TÜBİTAK-BİLGEM) since 2002. He has been working as project manager in several secure voice and data communication projects. He also continues his academic studies in collaboration

Comparative Analysis of CNN Models and Bayesian Optimization-Based Machine Learning Algorithms in Leaf Type Classification

Muhammet Fatih Aslan


Abstract— In this study, the leaves are classified by various Machine Learning (ML) and Deep Learning (DL) based Convolutional Neural Networks (CNN) methods. In the proposed method, first, image pre-processing is performed to increase the accuracy of the posterior process. The obtained image is a grayscale image without noise as a result of the pre-processing. These pre-processed images are used in classification with ML and DL. The Speeded Up Robust Features (SURF) are extracted from the grayscale image for ML-based learning. The features are restructured as visual words using the Bag of Visual Words (BoVW) method. Then, histograms are generated for each image according to the frequency of the visual word. Those histograms represent the new feature data. The histogram features are classified by four different ML methods, Decision Tree (DT), k-Nearest Neighbor (KNN), Naive Bayes (NB) and Support Vector Machine (SVM). Before using the ML methods, the Bayesian Optimization (BO) method, which is one of the Hyperparameter Optimization (HO) algorithms, is applied to determine hyperparameters. In the classification process performed with four different ML algorithms, the best accuracy is achieved with the KNN algorithm at 98.09%. Resnet18, ResNet50, MobileNet, GoogLeNet, DenseNet, which are state-of-the-art CNN architectures, are used for DL-based learning. CNN models have higher accuracy than ML algorithms.

Index Terms—Bag of Visual Words, Bayesian Optimization, Convolutional Neural Networks, Deep Learning, Speeded Up Robust Features

I. INTRODUCTION

IN AGRICULTURE, plants must be constantly observed to ensure the continuity of food production, and efforts must be made to ensure that different plant species do not become extinct. Plants are the essential source of life on earth. All living organisms on earth need nutrition and water. Plants release oxygen to the atmosphere through photosynthesis to meet these main needs. In addition, climate and the distribution of gases in the atmosphere become regular due to the plants.

MUHAMMET FATİH ASLAN, is with Department of Electrical Engineering, Karamanoglu Mehmetbey University, Karaman, Turkey, (e-mail: mfatihhaslan@kmu.edu.tr).

 <https://orcid.org/0000-0001-7549-0137>

Manuscript received Sep 12, 2022; accepted Nov 15, 2022.
DOI: [10.17694/bajece.1174242](https://doi.org/10.17694/bajece.1174242)

Many animals in nature meet their sheltering needs in the favor of plants. Substantial needs of people and animal such as drugs, fuel, etc. are provided by plants [1]. and animals in nature need plants, hundreds of plants are extinct due to human activities over the past 200 years. The continuity of plant production has an enormous effect on the future of human generation. Therefore, various precautions should be taken to prevent the extinction of the plant. For this purpose, one of the precautions taken is the identification of plant species [2].

Recognition is the most effective one for the protection of plant species. If plant species can be recognized accurately, precautions can be taken for the endangered plant species so that, the generation of plant species can be protected. But given the existence of nearly 3 million named and classified plant species on earth, this is no easy task[3]. Because it requires in-depth knowledge and experience about botany and plant systematics [4]. In addition, classification techniques such as morphological anatomy, cell biology, molecular biology, and photochemistry are complex and challenging topics, and therefore not suitable for online applications [5, 6]. As a result, the detection of all plant species on earth is an important but difficult task.

For easier plant identification, using plant leaves has long been accepted by researchers [7]. The leaves on the plant are the strongest determinant of that plant type. Because unlike flowers and fruits which appear in a short period, they are both outnumbered and long-lasting [8]. The exact color of the leaves can differ depending on the climate. That's the reason why it is essential to make a classification according to both shape and tissue features. In this way, more reliable results can be achieved [9]. In addition to recognizing the plant species, leaves also provide the discovery of important information such as plant development and plant disease. However, given the large number of leaf species, leaf recognition by traditional methods is a time-consuming and difficult task for botanists. Recently, such classification tasks have become easier and faster to be solved automatically with computer-aided systems [10].

Due to the disadvantages of traditional methods, studies using automatic identification methods have increased. Gaston and O'Neill [11] stated that it is possible to identify the plants automatically with the recent developments in the field of Artificial Intelligence (AI) and Digital Image Processing. The classification studies of the plants gained popularity in the past decade with developments in both Machine Learning (ML) and Deep Learning (DL) fields. For example, in papers by Kulkarni,

Rai, Jahagirdar and Upparamani [12], Kumar, Belhumeur, Biswas, Jacobs, Kress, Lopez and Soares [13] and Wäldchen, Rzanny, Seeland and Mäder [14] different computer vision and AI-based applications were introduced to improve leaf identification performance.

Plant identification studies provide significant benefits in applications for smart and precision agriculture. The development of automatic and rapid plant recognition applications can be used to know which plants have benefits and harms. In addition, leaf-based plant classification applications allow exciting applications for autonomous agricultural applications [15]. In this sense, the prominent application area is selective spraying. Vehicles that perform autonomous spraying according to the map of weed and crop distributions with selective spraying are an active field of application [16]. In addition, autonomous applications can be developed to monitor different crops grown in a greenhouse. Recently, different applications have been developed for precision agriculture with Unmanned Aerial Vehicles (UAV) and Unmanned Ground Vehicles (UGV). A comprehensive survey study on this was presented by Aslan, Durdu, Sabanci, Ropelewska and Gültekin [17]. As a result, plant identification applications are important and necessary in different fields for precision agriculture.

In this study, learning-based ML and DL methods are applied to identify leaves with high accuracy. For this, leaf images in the Folio [18] dataset are used. First, simple image processing steps are applied to raw images. In this way, irrelevant features that will negatively affect the artificial intelligence result are removed. ML and DL are implemented on these pre-processed images. While the feature extraction step is required first for traditional ML methods, feature extraction is performed automatically through deep layers in DL. In this study, to extract the necessary features for ML, Speeded Up Robust Features (SURF), a very powerful feature descriptor, is used. In addition, Bag of Visual Words (BoVW) is used to cluster a large number of features and express these clusters with histograms. Using the obtained histograms, the classification of leaf images is completed. In experimental studies, various ML methods such as Decision Tree (DT), k-Nearest Neighbor (KNN), Naive Bayes (NB) and Support Vector Machine (SVM) are applied with Bayesian Optimization (BO) to get the best performance from ML techniques. After the ML step, pre-processed leaves are classified with CNN models that can automatically extract features and perform classification. For this, frequently preferred methods such as Resnet18 [19], ResNet50 [19], MobileNet [20, 21], GoogLeNet [22] and DenseNet [23] are used. In the last step, the model that gives the most accurate result for the Folio dataset is also applied to the Swedish dataset to prove the result.

The hypothesis and novelty in this study are that traditional ML algorithms powered by SURF, BoVW and BO steps can perform close to popular CNN models in leaf recognition. The contributions of this article can be summarized as follows:

- 1) SURF features extracted from leaf images are strengthened with the BoVW method.
- 2) Different ML algorithms with powerful feature extraction and optimization steps are compared with different CNN models.

- 3) This study compares the high-level feature extraction power of CNN with ML methods including SURF, BoVW and BO.

- 4) At the end of the study, high leaf recognition successes are obtained.

The article is organized as follows. After the introduction, previous studies on leaf classification are discussed in section 2. Section 3 introduces leaf datasets used in learning algorithms. The pre-processing steps applied to leaf images, the BoVW approach, the ML and DL algorithms used, etc. are explained in section 4. Section 5 analyzes the ML and DL results obtained as a result of applying the proposed methods. Finally, section 6 concludes the work.

II. RELATED WORKS

Unfortunately, existing methods are not capable enough for the strong classification of the leaves [5, 24]. Therefore, many researchers aim to increase the accuracy of the classification by using different algorithms and their combinations. Vijaya Lakshmi and Mohan [25] suggested a new circle-based radii model. This model was based on the leaf's center point and leaf boundaries. 50 out of 220 classes from the ICL dataset were used. 44 features were extracted from each leaf. The SVM algorithm was used for classification. As a result of the study, a success of 93.33% was achieved. Koklu et al. [26] created grapevine leaf images and classified these images with the MobileNetv2 CNN model. They proposed three different approaches to classification. They implemented fine-tuned MobileNetv2 in the first, the MobileNetv2-SVM structure in the second, and the MobileNetv2-SVM-feature selection in the last. Also, different kernel functions for SVM were taken into account. At the end of the study, the CNN-SVM structure based on the selected features performed in the last step performed more successful classification with 97.60% accuracy. Sharma et al. [27] performed leaf classification based on color and texture features. They used HSV (hue, saturation, value) color space to extract color features and the Gray-Level Co-Occurrence Matrix (GLCM) algorithm to extract texture features. Extracted features were given as input to Artificial Neural Network (ANN), both individually and by fusion. At the end of the study, it was stated that the fusion features were stronger for leaf recognition. Arun and Viknesh [28] performed leaf classification on the Mendeley dataset using eleven different pre-trained CNN models such as AlexNet, Xception, ResNet50, and EfficientNet. In the comparisons made in the study, it was stated that the EfficientNet B5 model outperformed other models with 99.75% accuracy. Jin, Hou, Li and Zhou [24] emphasized leaf tooth features. The deformities in the leaf squares were removed with the Pauta criterion. After this, four-leaf tooth features (Leaf-obliqueness, Leaf-rate, Leaf-num and Leaf-sharpness, and Leaf-obliqueness) were obtained. As a result of the study, accuracy values between 73% and 80% were obtained. Saleem et al. [29], focused on handcrafted visual leaf features, feature extraction techniques, and classification methods for the analysis of visual leaf shape features in plant classification. In the experiments performed on the Flavia dataset, they presented a five-step algorithm for leaf recognition consisting of pre-processing and segmentation, feature extraction, etc. They used KNN, DT, NB, and multi-SVM

classifiers in the proposed algorithm. At the end of the study, KNN was the most successful ML method with 98.75% accuracy. Also, at the end of the study, the proposed method was compared with AlexNet. The proposed method was more successful than AlexNet in the presence of less training data. Dudi and Rajesh [30] performed a DL-based plant identification application using the Swedish and Mendeley dataset. First, they applied pre-processing steps such as median filtering and histogram equalization to the leaves. These images were then classified by CNN optimized with the Shark Smell-based Whale Optimization Algorithm (SS-WOA). They also introduced a threshold dependent classification method to classify untrained images. The proposed SS-WOA application outperformed different ML methods such as NB, SVM for both trained and untrained data. Sladojevic et al. [31] applied DL that was recently overemphasized in the field of ML. In that study, recognition of plant disease was aimed. Hence, leaf recognition was done using Convolutional Neural Networks (CNN). With this method, 13 different healthy leaf species were introduced to CNN. So, diseased species were distinguished. As a result of the study, the accuracy rate between 91% and 98% was acquired. Kulkarni et al. [12] utilized Zernike moment features in addition to shape, color, vein, and texture features. As a classifier, Radial Basis Probabilistic Neural Network (RBPNN) was preferred. The accuracy rate of the study using Flavia data was 93.82%. Wagle et al. [32] designed a compact CNN with different convolutional layers, named N1, N2, and N3, for plant species classification. They tested this compact CNN model on the PlantVillage and Flavia datasets. They then compared their results with the AlexNet model with transfer learning. The designed CNN model was superior in terms of training time and accuracy. The classification accuracies achieved were above 99%. Bakhshipour and Jafari [33] presented a proposal for the separation of sugar beet and weed. In the feature extraction phase, shape factors, moment invariants, and Fourier descriptors were extracted from images. ANN and SVM were carried out in the classification phase. In that study, four common weed species were detected in sugar beet. Accuracy rates were 92.92% for ANN and 95.00% for SVM. Finally, in another study by Kho et al. [34], three different *Ficus* species leaves (*F. benjamina*, *F. pellucidopunctata*, and *F. sumatrana*) were classified. Morphological features, Hu moment features, texture features, and Histogram of Oriented Gradients (HOG) were employed in the feature extraction stage. ANN and SVM were applied to perform classification. As a result of both models, the accuracy rate was found as 83.3%.

In general, previous studies have used different feature extraction, optimization and classification algorithms for leaf classification. In previous studies, it is desired to increase the accuracy of leaf recognition by using different methods together (hybrid). Some studies perform classification with ML, others with DL. The learning algorithm used in leaf classification is vital for accuracy. No comprehensive study involving DL and ML comparison has been carried out so far. The superiority of DL algorithms over ML is known. In this sense, this study equips ML algorithms in the best way and compares them with different CNN models.

III. DATASET

A. Folio Dataset

In this study, the dataset from UCI[35] library is used. The Folio dataset includes 32 different images each belonging to different leaves. Each type has 20 images. Images are obtained from a mobile phone with a 1980x1024 resolution. In this study, 15 different types of leaves from the Folio [18] dataset are classified. The types of leaves are shown in Fig. 1 (a).

B. Swedish Leaf Dataset

The Swedish Leaf dataset [36] is a public dataset created by Linköping University and the Swedish Museum of Natural History. The dataset contains a total of 15 different leaf types, and each type has 75 sample images. Some species are very similar in shape to each other. The image background is white. The leaf types of the Swedish Leaf dataset are shown in Fig. 1(b).



(b) Swedish Leaf dataset
Fig.1. Leaf datasets used in the study

IV. METHODOLOGY

The methodology part of this study deals with image processing, feature extraction, ML and DL methods. A general

flow chart showing the application of these methods in this study is shown in Fig. 2. In this section, each method used is examined in detail and the outputs of the relevant method are shared.

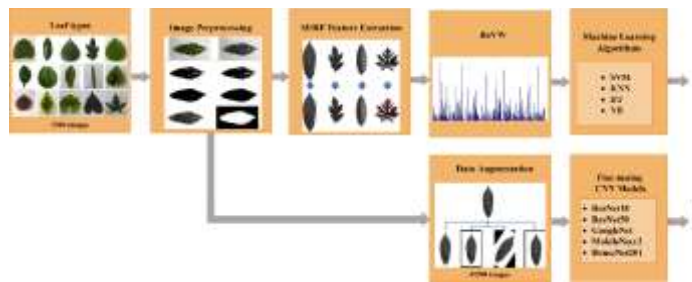


Fig.2. Methods used in the study

A. Image Processing

Most of the time, it is crucial to perform pre-processing operations in the original image to minimize the error of the feature extraction process. Using the advantage of pre-processing, some operations like the extraction of desired features, minimizing the noise can be achieved, and it directly affects the performance of the classification. This study presents a comparative analysis of ML and DL methods, which are frequently preferred in leaf classification applications. The main purpose of this study is to compare the results obtained from ML and DL-based algorithms. However, in order to increase the accuracy and reliability of both ML and DL-based algorithms, it is necessary to examine only leaf-containing pixels. In this sense, some image pre-processing steps are applied for raw leaf images. The image is belonging to the Croton plant, and each pre-processing step is illustrated in the order in Fig. 3.

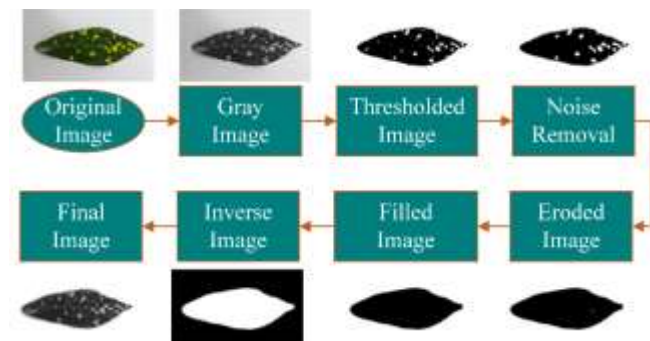


Fig.3. Pre-processing steps

As shown in Fig. 3, the original image is converted to a grayscale image to overcome the complexity derived from color and to minimize the noise. When Fig. 1 (a) is examined, it is obvious that backgrounds are completely white for some images while it is not for others. To improve recognition performance, features should be extracted mostly from the leaf. For this, a noise-free and uniform background should be provided for all leaves. For this reason, the background of all images should be white. The threshold value is determined by the Otsu method [37], and the leaf is completely separated from the background to cover the background with white color. The main purpose of the pre-processing is that the background should be white, and the leaves are gray-scale. For this reason,

noise removal, morphological operations, filling, and inversion were performed on the image after the determination of the threshold value of the image. Eventually, the multiplication of achieved inverse image and gray image gives the final image. After this stage, most of the features extracted from the final image belong to the leaf. In summary, as a result of the pre-processing, the original image is subjected to specific processing, and thus, noise is removed, and gray space leaf images are obtained. Each of these processes is first applied to all leaves. The obtained results are recorded, and the features of these images are extracted. The original images and the final images of the Croton plant as a result of pre-processing are shown in Fig. 4.

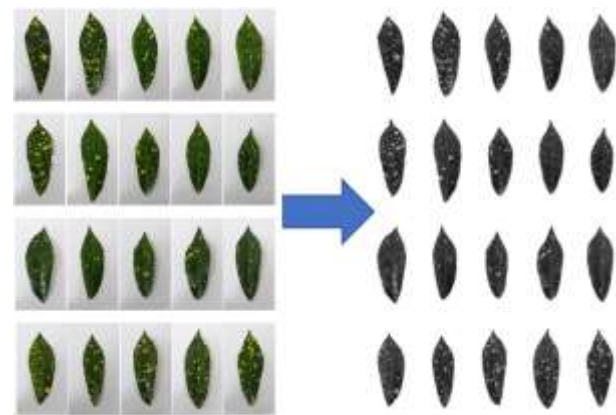


Fig.4. Original and final images of Croton plant

B. K-Means Algorithm

The K-Means algorithm is a frequently used learning method that has a simple structure. In this method, a large number of extracted features are divided into k groups. When the algorithm starts, first of all, centers are determined. Afterwards, data are classified according to these centers. According to the classification results, the center points are constantly updated. This process continues iteratively. Clustering is completed when the error between the specified center point and the data reaches the minimum value [38, 39].

C. Bag of Visual Words

The BoVW algorithm is an adapted version of the BoW (Bag of Words) algorithm to images. Using the BoW algorithm, it is possible to obtain a histogram of the words in a document. In this way, document classification can be performed. Similarly, with the BoVW, images can be classified by identifying visual words in an image and creating histograms [40]. In the BoVW method, the extracted features are clustered first. As a result of clustering, words are obtained, and each cluster represents a word. Finally, histograms are created according to the word frequency in the input image [41, 42].

D. Feature Extraction

In this study, SURF features are extracted from each pre-processed image. SURF is superior to SIFT (Scale Invariant Feature Transform) in terms of faster feature extraction [43]. Since there are 20 pieces from each leaf and the training rate is 80% in the classification, a total of 240 leaves are used to

extract features. The total number of extracted features is 947085. In Fig. 5, the features of some different leaves are shown with red marks.

ML or DL models can be fed directly with SURF features. However, more successful results were obtained in studies [44-47] that applied SURF features together with the BoVW model. For the detection of SURF key points, the Hessian matrix is used. With the determinant of the Hessian matrix, it is decided whether a point will be chosen as the key point. Considering a pixel, its Hessian matrix is calculated as follows:

$$H(f(x, y)) = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix} \quad (1)$$

In an I image, given the point $\mathbf{x} = [x, y]$, the Hessian matrix $H(\mathbf{x}, \sigma)$ at point \mathbf{x} on the scale σ is defined as follows [48]:

$$H(\mathbf{x}, \sigma) = \begin{bmatrix} L_{xx} & L_{xy} \\ L_{yx} & L_{yy} \end{bmatrix} \quad (2)$$

Here $L_{xx}(\mathbf{x}, \sigma)$ is the convolution of the image I at the \mathbf{x} point and the Gaussian quadratic derivative ($\frac{\partial^2}{\partial x^2} G(x, y, \sigma)$). Others (L_{xy} , L_{yx} , L_{yy}) are calculated similarly. Since key point calculations will be made for each pixel, the SURF method is affected by noise in the image and faulty or missing features can be extracted [49-51]. Because SURF features are calculated for each pixel. Therefore, noise in the image and unnecessary blobs in the background should be removed. Various noise removal or denoising methods can be used for this purpose [52]. By removing the noises, only features are extracted from the main object. The number of features extracted with SURF may be different for each image. In other words, the number of pixels that are resistant to rotation and scale change in an image can be different. However, thanks to the use of these features with the BoVW model, the number of features will equate to the number of visual words.

As seen in Fig. 5, the background noise is completely removed, and the SURF features are only on the leaf. These SURF features are then processed in the BoVW algorithm to generate new feature data. In this way, it is aimed to represent more complex, and numerous SURF features in a more meaningful way with BoVW. Features are first clustered with BoVW. In this study, k-means is used as the clustering algorithm, and k value is set to 500. In other words, 947085 features are distributed to 500 clusters in k-means. Each cluster represents a word. The result of BoVW is expressed using histograms according to an image word frequency. However, the histogram only gives the word frequency in an image. Location information is not available on the histogram. The histogram obtained from a sample leaf image is shown in Fig. 6.

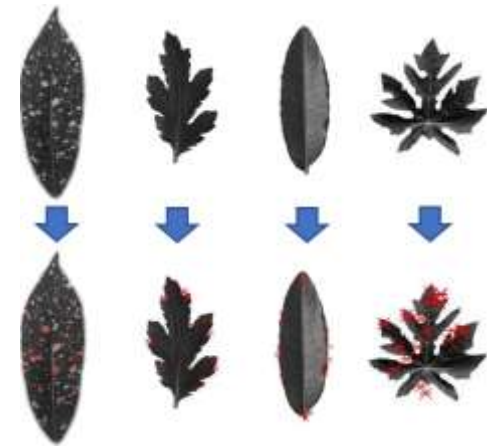


Fig.5. SURF features extracted from pre-processed images

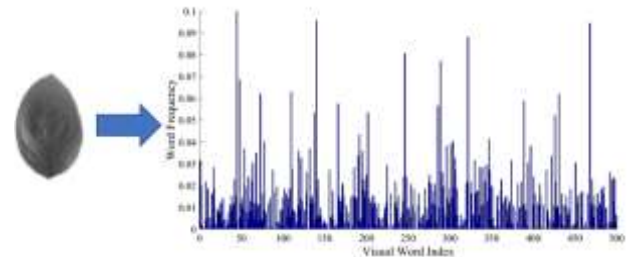


Fig.6. Visual word frequency histogram obtained from a leaf of Ashanti Blood

The histogram obtaining process in Fig. 6 is carried out for each leaf. As a result, a new numerical dataset in the size of 240x500 is generated for 240 images. These data can now be classified by four different ML methods.

E. Machine Learning-Based Leaf Classification

In this study, four different ML methods, DT, KNN, SVM, and NB, respectively are used for the classification of the leaves. In this part, methods are explained briefly.

1) Decision Tree

DT is a supervised learning method that can solve classification and regression problems by using trees, leaves, and branches in a representative manner. The tree structure is formed with DT according to the used dataset. A tree consists of leaves and branches. Leaves represent the label of the class while branches represent the cluster that forms the label class. Forming a tree continues iteratively until there are no criteria to divide the samples. After the forming phase of the tree, it is time for the pruning phase which increases the performance of the classification. A big dataset can be classified successfully by using DT [53-55].

2) k-Nearest Neighbors

The KNN algorithm is an unsupervised learning method that is easy to understand and implement. The algorithm works by considering the distance between the data of known class and a new sample for the classification operation. The number k represents the number of data closest to the new sample, and the class assignment is performed by considering the majority. Contrary to other methods, there is no time lost for the training process, in another saying there is no training process. Although KNN is a classical method, it is still preferred because of its

accuracy and speed, and it was used frequently in applications such as pattern recognition and classification [56-58].

3) Support Vector Machine

SVM is a supervised learning method for classification and regression applications because of its high performance. The training process was performed in a multi-dimensional space by creating hyperplanes. The classes are separated with hyperplanes, and for a strong separation, hyperplanes should divide the class with an optimal distance. SVM includes the necessary calculations that form the hyperplanes to provide optimum separation. In addition, a multi-dimensional and non-linear dataset can also be classified with SVM. The kernel trick method is used for this purpose. In this way, the feature space is mapped to a higher space, and afterwards, hyperplanes are created [59-61].

4) Naive Bayes

NB is a probabilistic classification model based on the Bayes theorem. With probability calculations using data of a known class, the class of the new data can be calculated. In this method, classification is made by assuming that the features have connections and relations. One of the most important advantages of NB is that it can perform classification with little training data. Moreover, NB is less affected by noises [62-64].

F. Bayesian Optimization (BO)

BO [65] is a hyperparameter search algorithm. It was developed using Bayes' theorem. With Bayes, the posterior distribution is estimated using prior knowledge. The equation representing the general Bayes theorem is shown in Equation 3. According to Equation 3, posterior distribution ($P(X|Z)$) is directly proportional to likelihood ($P(Z|X)$) and a priori distribution (bias) ($P(X)$) [66].

$$P(X|Z) \propto P(Z|X)P(X) \quad (3)$$

$$x^* = \arg \min_f(x) \quad x \in X \quad (4)$$

BO is often used among HO methods. The general calculation formula for this optimization method is as in Equation 4. In Equation 4, the value x^* minimizes the function $f(x)$. Here x represents a hyperparameter of the ML algorithm. $f(x)$ represents an objective function, also called the Gaussian process model, that strives to be minimized. The hyperparameter type x is searched in the X search space. Finding the optimal hyperparameter for ML algorithms in the search space is quite costly. Also, it would be costly to use their values in the whole space to find the optimum value of x as in Equation 4. An iterative approach is required to reduce cost. At this stage, BO converges to the optimum value by applying Bayes-based iterative optimization[67].

G. Deep Learning-Based Leaf Classification

This section presents the data augmentation step applied for leaf classification based on DL. It also briefly introduces the Resnet18 [19], ResNet50 [19], MobileNet [20, 21], GoogLeNet [22] and DenseNet [23] CNN models applied for classification.

1) Data Augmentation

It is a well-known fact that a rich dataset increases the success of DL classification. However, it may not always be

possible to obtain a rich dataset. Therefore, the images in the dataset can be artificially augmented to obtain a rich dataset using various techniques [68, 69]. Data augmentation is actions performed on raw images to increase the amount of data in the dataset.

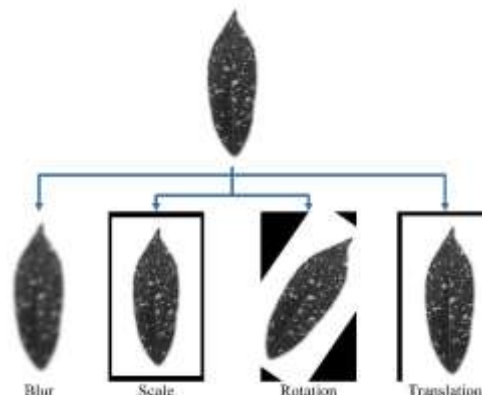


Fig. 7. Data augmentation techniques and sample result images

In this study, there are a total of 300 images, 20 of which belong to each of 15 different leaf types. Since this number of images is insufficient for a DL study, data augmentation techniques are applied. Four different data augmentation techniques (Rotation, Scale, Translation, and Blur) are applied to each image shown in Fig. 7 to ensure data diversity. With the rotation operation, the image is rotated around its center by a certain degree. The translation process is to shift the images right, left, up or down on a pixel basis. With scale operation, the scale of the image is changed at a certain interval, so that the image of the object in different dimensions is created. For blur, a low-pass filter is applied to reduce the high-frequency effect in the image. The total number of images is increased to 1500 by using data augmentation techniques. The lower and upper limit values of these data augmentation methods are shown in Table I.

TABLE I
THE LOWER AND UPPER LIMITS OF THE DATA AUGMENTATION TECHNIQUES USED

Parameter	Lower Limit	Upper Limit
Blur	1	4
Rotation	-45°	45°
Scale	0.8	1.2
Translation (pixel)	-20	+20

2) ResNet18 and ResNet50

CNN architectures are designed in a deeper structure over time. This complicates these architectures. Also, constantly increasing the depth does not always increase success. ResNet [19] is developed to avoid the vanishing gradient problem, which limits the learning of the network as the depth increases. To achieve this, residual blocks are used to transmit residual values to subsequent layers. ResNet18 and ResNet50 models with different depths such as 18 and 50 were developed using this principle.

3) MobileNet

MobileNet is a deep architecture with low computation and fast processing capability, designed for use on mobile devices. Expensive convolution layers are replaced with 1x1 and 3x3

convolution layers to reduce the number of parameters. In this way, fewer learning parameters are obtained. MobileNetv2, on the other hand, uses inverted residual blocks with bottleneck properties and thus has fewer parameters [70].

4) GoogLeNet

GoogLeNet is a 22-layer architecture developed by researchers at Google. It is the first version of the Inception network (Inception v1). In its architecture, it used an inception model consisting of convolution filters with dimensions of 1x1, 3x3 and 5x5. In this way, while increasing the depth of the network, the computational load is reduced [22].

5) DenseNet

It is similar to the ResNet architecture in general, but denser. Each convolution layer except the first layer receives feature maps of all previous convolution layers. That is, the feature maps of a convolution layer are given as input to all subsequent layers. DenseNet-201 is a deep CNN with a layer depth of 201 [23].

V. RESULTS

This section evaluates the classification performance of ML and DL algorithms. Performance metrics used for this are given in Equations 5-10.

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN} \quad (5)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (6)$$

$$\text{Sensitivity} = \frac{TP}{TP + FN} \quad (7)$$

$$\text{Specificity} = \frac{TN}{TN + FP} \quad (8)$$

$$\text{F1 - score} = \frac{2TP}{2TP + FP + FN} \quad (9)$$

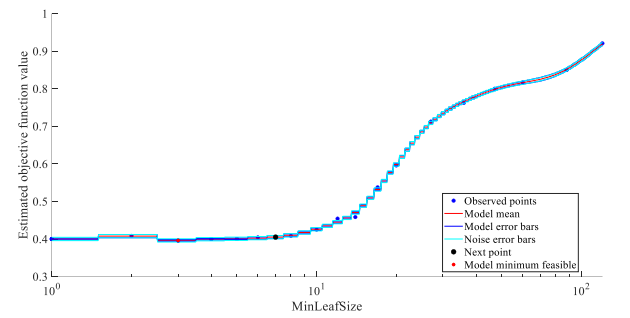
TP: True Positive TN: True Negative FP: False Positive FN: False Negative

A. ML Results

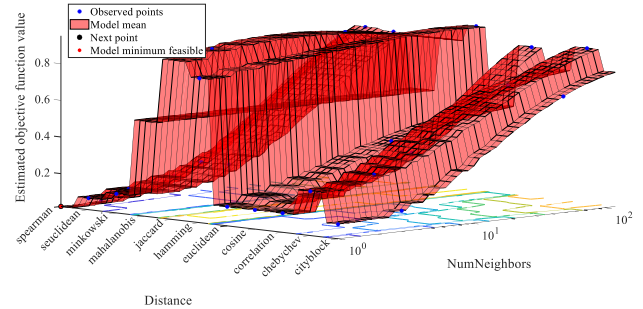
DT, KNN, SVM, and NB have different hyperparameters that significantly affect the accuracy of the algorithm. When these hyperparameters are determined, optimal success can be obtained from the relevant ML method. With the Hyperparameter Optimization (HO) method, hyperparameter values that make system error minimum are determined [71]. BO [65] approach which is one of the HO methods is used in this study to obtain parameters that would give maximum success rates. For this reason, hyperparameters are first determined by the BO approach and then classification is made by using these hyperparameters. The hyperparameter values used for DT, KNN, NB, and SVM and the time taken to find the optimum hyperparameters (BO duration) are shown in Table II.

In addition, optimization graphs determining DT, KNN, and NB hyperparameters are shown in Fig. 8. Fig. 8 shows the objective function values according to the changes of the

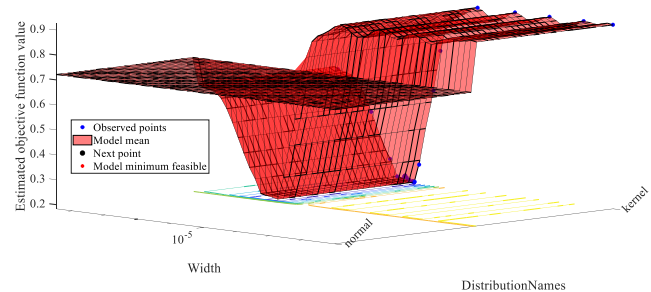
hyperparameters in three different ML algorithms. In these graphs, the values that minimize the objective function are searched in the search space. The hyperparameter value of each ML algorithm is calculated iteratively with BO and the objective function is recalculated accordingly. Instead of calculating the objective function with all space values, BO reaches the optimum value faster with an iterative approach. For each ML algorithm, the hyper parameter values that minimize the graph are recorded and ML classifications are made using these values. For SVM, the four-dimensional BO graph could not be plotted due to the number of variables (objective function, C, kernel scale, and coding method). However, results are obtained. The confusion matrices obtained after the classification using the specified parameter value are shown in Fig. 9, respectively. The metric values in Equations 1-6 are calculated to determine the classification performance. These values are shown in Table III.



(a) DT



(b) KNN



(c) NB

Fig. 8. BO results of ML algorithms

Considering the confusion matrices in Fig. 9 and the performance metrics in Table III, it is seen that KNN and SVM methods are more successful than other methods with an

accuracy of 98.09%. Since KNN is the ML method that provides faster training, KNN is a more preferable method in this respect. It should be noted that these results are calculated using the ML parameters obtained with BO, that is, they are the optimum classification results.

TABLE II
HYPERPARAMETERS FOR ML ALG. AND BO DURATION VALUES

ML Alg.	DT	KNN	NB	SVM
Hyperparameters	Min. leaf size: 3	Number of neighbors: 1 Distance type: spearman	Distribution type: kernel Kernel width: 0.00308	Coding method: one-versus-all Box Constraints (C): 0.84721 Kernel scale: 0.0336963
BO duration (sec.)	61.86	41.944	3664.401	1616.419

TABLE III
PERFORMANCE METRICS OF ML ALGORITHMS

ML algorithm	Acc.	Spec.	Prec.	Sens.	F1-Score
DT	0.6536	0.9752	0.7213	0.6536	0.6536
KNN	0.9809	0.9986	0.9851	0.9809	0.9805
NB	0.8963	0.9926	0.9173	0.89635	0.8958
SVM	0.9809	0.9986	0.9851	0.9809	0.9805

B. DL Results

In the next step after data augmentation, five different popular CNN models are fed with augmented leaf images. These are ResNet18, ResNet50, GoogLeNet, DenseNet201, and MobileNetv2, which are frequently used in DL studies. The

architecture of each model consists of different numbers and types of layers. The input image size for each model is 224×224 . Therefore, each image is resized and given to the models. Of 1500 leaf images, 80% were used for training and 20% for testing.

Before training, the dataset is divided into small groups. Learning is done in these small groups. This parameter, called mini-batch size, determines how many data the CNN model can process at the same time. The groups separated by mini-batch are trained sequentially and the weights of the network are updated. The number of times that all training data is given to the network is called the epoch. In this way, the network is repeatedly trained and the weights are updated to reduce the error. The update rate of the weights is determined by the Learning Rate. The parameter values used for training in all CNN models are as follows: Execution Environment: GPU, Max. Epoch: 20, Learning Rate: 0.001, and Mini Batch Size: 32. The optimization algorithm used to train and reduce the loss value is Stochastic Gradient Descent with Momentum (SGDM). Result values are obtained after training and testing for each CNN model. The training of CNN models is carried out on a laptop computer with Intel Core i7-7700HG CPU, NVIDIA GeForce GTX 1050 4 GB graphics card and 16 GB RAM. The confusion matrices obtained according to the ResNet18, ResNet50, GoogLeNet, DenseNet201, and MobileNetv2 classification results are shown in Fig. 10. In addition, accuracy, specificity, precision, F1-score, and sensitivity performance metrics were calculated according to these error matrices and these values are given in Table IV.

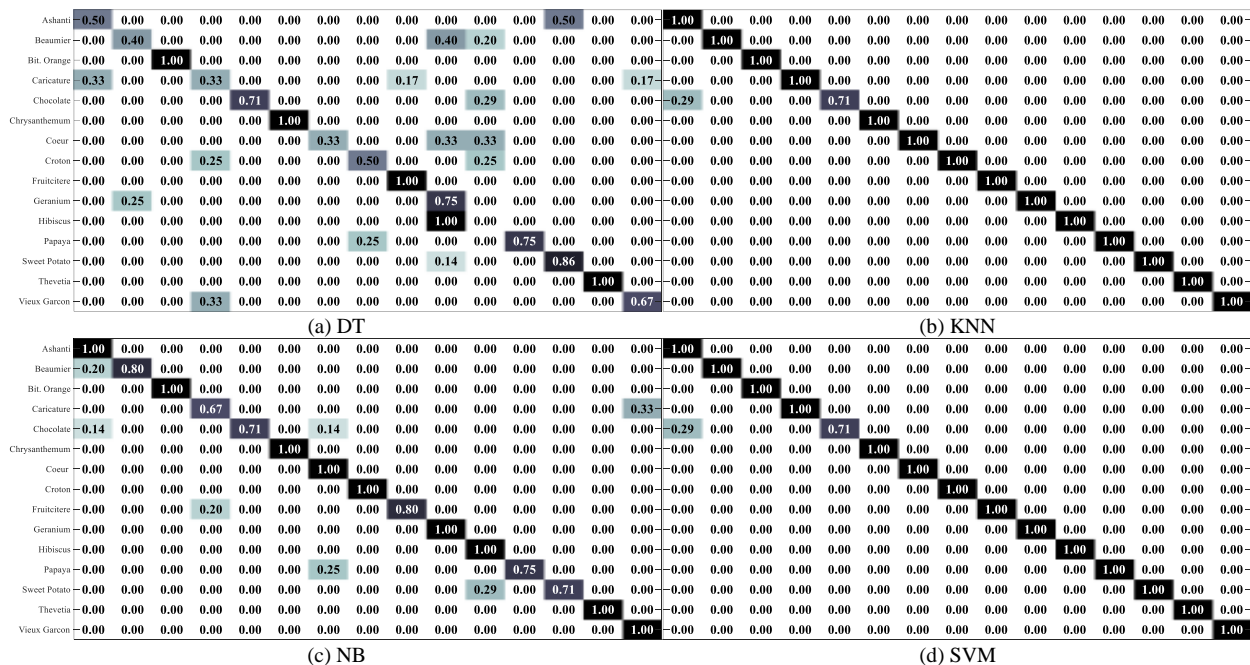


Fig. 9. Confusion matrices of ML algorithms

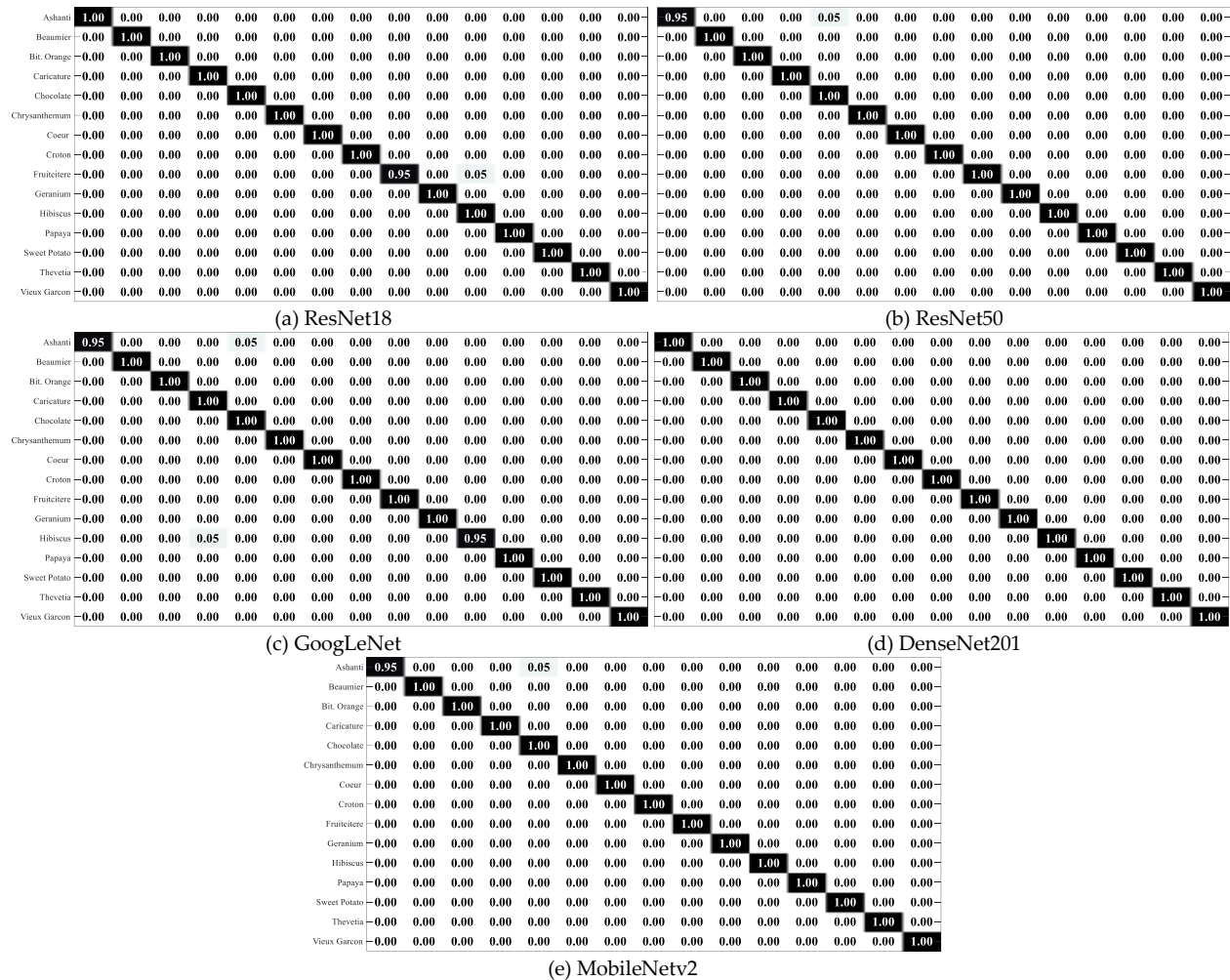


Fig. 10. Confusion matrices of DL algorithms

TABLE IV
PERFORMANCE METRICS OF CNN MODELS

Model	Acc.	Spec.	Prec.	Sens.	F1-Score
ResNet18	0.9967	0.9998	0.9970	0.9968	0.9968
ResNet50	0.9967	0.9998	0.9974	0.9970	0.9971
GoogLeNet	0.9933	0.9995	0.9923	0.9938	0.9928
DenseNet201	1	1	1	1	1
MobileNetV2	0.9967	0.9998	0.9974	0.9970	0.9971

When the confusion matrices in Fig.10 and the classification results in Table IV are examined, it is seen that all the CNN models used provide superior success. DenseNet201, which has the densest connectivity among CNN models, correctly classified all leaves.

In addition, in this study, DL and ML algorithms, which provide the highest classification (SVM for ML, DenseNet201 for DL) are also tested on the Swedish leaf dataset. First, each of the leaf images in the Swedish dataset (see Fig.1(b)) is pre-

processed as in the Folio dataset. 80% of these images are used for training and the rest for testing. For ML, SURF features are extracted from the processed images. Then these features are processed with the BoVW method and new word features are created. Afterwards, the BO-optimized SVM method classifies leaf images with 96.89% accuracy. The confusion matrix obtained as a result of classification with SVM is shown in Fig. 11(a). Other metric values are also shown in Table V.

For DL, the number of pre-processed images is increased by the data augmentation techniques shown in Fig. 7. As a result of increasing the data, the total number of processed leaf images increases to 5625. After the training and testing steps, the confusion matrix obtained with the DenseNet201 network for the test data is shown in Fig. 11(b). In addition, other performance metrics are shared in Table V. Accordingly, Swedish leaf dataset images are classified with 99.91% accuracy with DenseNet201. The resulting accuracy and other metrics prove the robustness of the proposed method and support the results obtained on the Folio dataset.

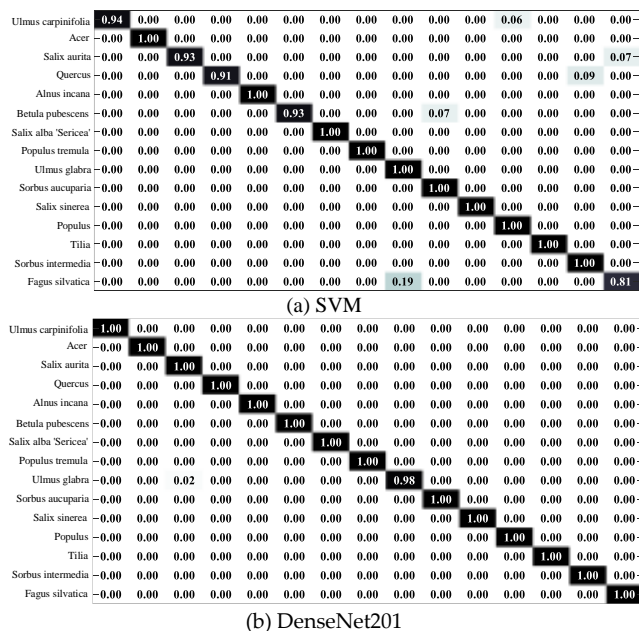


Fig. 11. Confusion matrices obtained with DenseNet201 and SVM for the Swedish Leaf dataset

TABLE V

PERFORMANCE METRICS OF SVM AND DENSENET201 MODEL FOR THE SWEDISH LEAF DATASET

Model	Acc.	Spec.	Prec.	Sens.	F1-Score
SVM	0.9689	0.9978	0.9706	0.9685	0.9680
DenseNet201	0.9991	0.9999	0.9992	0.9988	0.9990

VI. CONCLUSION, DISCUSSION AND FUTURE WORKS

In this study, leaf types are classified using ML and DL-based approaches. For a more comprehensive classification, four different ML methods, namely DT, KNN, SVM, and NB, and five different DL methods, ResNet18, ResNet50, GoogLeNet, DenseNet201, and MobileNetv2, are applied. Folio dataset images are used in all these learning-based algorithms. Then the Swedish leaf dataset is also used to prove the robustness of the proposed method. No features are directly extracted from the raw leaf images in this dataset. Firstly, image pre-processing steps are applied to the raw leaf images. Noise and background are removed during pre-processing so that only features are extracted based on the leaves. In this way, the basis for a more successful classification is formed. Pre-processed images are saved, and classifications are performed using these images.

In ML-based classification, the SURF features are extracted from the saved images. Thus, scale-invariant and rotation-invariant features are obtained. However, there are many feature points around 947085 with SURF. BoVW is used to classify these features in a shorter time and to reduce their size by combining them to be more meaningful. SURF features are clustered with BoVW and then histograms are generated according to the number of clusters in each image. That means each image is expressed by the histogram of visual words. K-means is used for clustering. As a result, the histograms are classified by four different ML methods (DT, KNN, SVM, NB). BO method, which is one of the HO algorithms, is used with ML methods to obtain the best results. In this way, optimum

hyperparameters are determined for each ML method. As a result, for DT, KNN, SVM, and NB, the classification accuracy rates are 65.36%, 98.09%, 98.09%, and 89.63%, respectively. KNN is a faster learning algorithm than SVM. So, the highest performance in terms of both accuracy and time is provided by KNN.

In the next stage, the types of leaves are classified with different state-of-the-art CNN models. New leaf images have been obtained using data augmentation techniques. In this way, the number of data important for deep network training is increased. Finally, the augmented images were classified using five different CNN models: ResNet18, ResNet50, GoogLeNet, DenseNet201, and MobileNetv2. The classification accuracy rates of these CNN models are 99.67%, 99.67%, 99.33%, 100%, and 99.67%, respectively. In addition, at the end of the study, ML and DL models that successfully classify the Folio dataset are analyzed on the Swedish Leaf dataset. As a result, leaf images in the Swedish Leaf dataset are classified with 99.91% accuracy with DenseNet201 and 96.89% with SVM.

The results show that CNN-based methods for leaf classification are more successful than traditional image processing and ML methods. Finally, it is considered that these results can be useful and encouraging for future studies. As more complex features are obtained from more images with DL, higher accuracy is expected. However, the success of KNN and SVM methods is quite high despite fewer features and fewer images. The combination of SURF, BOVW, and BO methods provided a powerful feature map for ML algorithms.

As a result of both ML and DL experiments, a successful leaf recognition application was carried out for pre-processed leaf images. The highest leaf recognition accuracy was 98.09% with ML and 100% with DL. These accuracies were obtained with KNN/SVM and DenseNet201 models, respectively. The performance of ML methods powered by SURF, BoVW, and BO close to CNN models supports the initial hypothesis.

Despite these successful results, ML and DL applications have some limitations. For a real-time application, extraction of SURF features and BO would cause significant delays and thus ML implementations would be insufficient for practical implementation. Both ML and DL applications should be able to recognize faster. Although data augmentation significantly impacts the success of DL, these images are artificial. Therefore, performing DL experiments on a larger dataset will provide more reliable results.

Considering the implementation and results of ML and DL, ML provided a more erroneous result than DL despite the SURF, BoVW, and BO implementation steps. However, data augmentation was performed for DL as an additional step. Ultimately, however, DL-based CNN models proved to be both an easier and more successful tool for leaf recognition. DL-based solutions will provide stronger recognition if larger datasets are used. However, increasing the accuracy does not only depend on the dataset. The design of the deep architecture for training is also important. For this reason, more successful results can be obtained with different deep architectures in future studies. In addition, CNN-SVM networks, which have come to the fore recently and which extract features with DL and perform classification with SVM, can provide stronger leaf recognition performance. Besides architecture, the fusion of features extracted from leaves in different ways can also enhance recognition performance. In this context, the fusion of DL and ML features may provide a stronger representation. In

addition to all these, the current success can be increased by applying the HO applied in ML algorithms to CNN models similarly.

REFERENCES

- [1] J. S. Cope, D. Corney, J. Y. Clark, P. Remagnino, and P. Wilkin, "Plant species identification using digital morphometrics: A review," *Expert Systems with Applications*, vol. 39, no. 8, pp. 7562-7573, 2012.
- [2] Z.-Q. Zhao, L.-H. Ma, Y.-m. Cheung, X. Wu, Y. Tang, and C. L. P. Chen, "ApLeaf: An efficient android-based plant leaf identification system," *Neurocomputing*, vol. 151, pp. 1112-1119, 2015.
- [3] B. Harish, A. Hedge, O. Venkatesh, D. Spoorthy, and D. Sushma, "Classification of plant leaves using Morphological features and Zernike moments," pp. 1827-1831.
- [4] C. Zhao, S. S. Chan, W.-K. Cham, and L. Chu, "Plant identification using leaf shapes—A pattern counting approach," *Pattern Recognition*, vol. 48, no. 10, pp. 3203-3215, 2015.
- [5] X. Wang, J. Liang, and F. Guo, "Feature extraction algorithm based on dual-scale decomposition and local binary descriptors for plant leaf recognition," *Digital Signal Processing*, vol. 34, pp. 101-107, 2014.
- [6] K. K. Thyagarajan, and I. Kiruba Raji, "A Review of Visual Descriptors and Classification Techniques Used in Leaf Species Identification," *Archives of Computational Methods in Engineering*, vol. 26, no. 4, pp. 933-960, 2019/09/01, 2019.
- [7] J. Chaki, R. Parekh, and S. Bhattacharya, "Plant leaf classification using multiple descriptors: A hierarchical approach," *Journal of King Saud University - Computer and Information Sciences*, vol. 32, no. 10, pp. 1158-1172, 2020/12/01, 2020.
- [8] A. Aakif, and M. F. Khan, "Automatic classification of plants based on their leaves," *Biosystems Engineering*, vol. 139, pp. 66-75, 2015.
- [9] L. Longlong, J. M. Garibaldi, and H. Dongjian, "Leaf classification using multiple feature analysis based on semi-supervised clustering," *Journal of Intelligent & Fuzzy Systems*, vol. 29, no. 4, pp. 1465-1477, 2015.
- [10] J. Su, M. Wang, Z. Wu, and Q. Chen, "Fast Plant Leaf Recognition Using Improved Multiscale Triangle Representation and KNN for Optimization," *IEEE Access*, vol. 8, pp. 208753-208766, 2020.
- [11] K. J. Gaston, and M. A. O'Neill, "Automated species identification: why not?," *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, vol. 359, no. 1444, pp. 655-667, 2004.
- [12] A. Kulkarni, H. Rai, K. Jahagirdar, and P. Upparamani, "A leaf recognition technique for plant classification using RBPNN and Zernike moments," *International Journal of Advanced Research in Computer and Communication Engineering*, vol. 2, no. 1, pp. 984-988, 2013.
- [13] N. Kumar, P. N. Belhumeur, A. Biswas, D. W. Jacobs, W. J. Kress, I. C. Lopez, and J. V. Soares, "Leafsnap: A computer vision system for automatic plant species identification," *Computer vision—ECCV 2012*, pp. 502-516: Springer, 2012.
- [14] J. Wäldchen, M. Rzanzy, M. Seeland, and P. Mäder, "Automated plant species identification—Trends and future directions," *PLoS computational biology*, vol. 14, no. 4, pp. e1005993, 2018.
- [15] D. Hall, C. McCool, F. Dayoub, N. Sunderhauf, and B. Upcroft, "Evaluation of features for leaf classification in challenging conditions," pp. 797-804.
- [16] B. Liu, and R. Bruch, "Weed detection for selective spraying: A review," *Current Robotics Reports*, vol. 1, no. 1, pp. 19-26, 2020.
- [17] M. F. Aslan, A. Durdu, K. Sabanci, E. Ropelewska, and S. S. Gültekin, "A Comprehensive Survey of the Recent Studies with UAV for Precision Agriculture in Open Fields and Greenhouses," *Applied Sciences*, vol. 12, no. 3, pp. 1047, 2022.
- [18] T. Munisami, M. Ramsurn, S. Kishnah, and S. Pudaruth, "Plant leaf recognition using shape features and colour histogram with K-nearest neighbour classifiers," *Procedia Computer Science*, vol. 58, pp. 740-747, 2015.
- [19] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," pp. 770-778.
- [20] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam, "Mobilenets: Efficient convolutional neural networks for mobile vision applications," *arXiv preprint arXiv:1704.04861*, 2017.
- [21] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "Mobilenetv2: Inverted residuals and linear bottlenecks," pp. 4510-4520.
- [22] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," pp. 1-9.
- [23] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," pp. 4700-4708.
- [24] T. Jin, X. Hou, P. Li, and F. Zhou, "A novel method of automatic plant species identification using sparse representation of leaf tooth features," *PLoS One*, vol. 10, no. 10, pp. e0139482, 2015.
- [25] B. Vijaya Lakshmi, and V. Mohan, "Plant leaf image detection method using a midpoint circle algorithm for shape-based feature extraction," *Journal of Modern Applied Statistical Methods*, vol. 16, no. 1, pp. 26, 2017.
- [26] M. Koklu, M. F. Unlarsen, I. A. Ozkan, M. F. Aslan, and K. Sabanci, "A CNN-SVM study based on selected deep features for grapevine leaves classification," *Measurement*, vol. 188, pp. 110425, 2022/01/01, 2022.
- [27] G. Sharma, A. Kumar, N. Gour, A. K. Saini, A. Upadhyay, and A. Kumar, "Cognitive framework and learning paradigms of plant leaf classification using artificial neural network and support vector machine," *Journal of Experimental & Theoretical Artificial Intelligence*, pp. 1-26, 2022.
- [28] Y. Arun, and G. S. Viknesh, "Leaf Classification for Plant Recognition Using EfficientNet Architecture," pp. 1-5.
- [29] G. Saleem, M. Akhtar, N. Ahmed, and W. S. Qureshi, "Automated analysis of visual leaf shape features for plant classification," *Computers and Electronics in Agriculture*, vol. 157, pp. 270-280, 2019/02/01, 2019.
- [30] B. Dudi, and V. Rajesh, "Optimized threshold-based convolutional neural network for plant leaf classification: a challenge towards untrained data," *Journal of Combinatorial Optimization*, vol. 43, no. 2, pp. 312-349, 2022/03/01, 2022.
- [31] S. Sladojevic, M. Arsenovic, A. Anderla, D. Culibrk, and D. Stefanovic, "Deep neural networks based recognition of plant diseases by leaf image classification," *Computational intelligence and neuroscience*, vol. 2016, 2016.
- [32] S. A. Wagle, R. Harikrishnan, S. H. M. Ali, and M. Faseehuddin, "Classification of Plant Leaves Using New Compact Convolutional Neural Network Models," *Plants*, vol. 11, no. 1, pp. 24, 2022.
- [33] A. Bakhshipour, and A. Jafari, "Evaluation of support vector machine and artificial neural networks in weed detection using shape features," *Computers and Electronics in Agriculture*, vol. 145, pp. 153-160, 2018.
- [34] S. J. Kho, S. Manickam, S. Malek, M. Mosleh, and S. K. Dhillon, "Automated plant identification using artificial neural network and support vector machine," *Frontiers in Life Science*, vol. 10, no. 1, pp. 98-107, 2017.
- [35] "UCI Machine Learning Repository," 07.05.2022, 2022; <https://archive.ics.uci.edu>.
- [36] "Swedish Leaf Dataset," 10.08.2022, 2022; <https://www.cvl.isy.liu.se/en/research/datasets/swedish-leaf/>.
- [37] N. Otsu, "A Threshold Selection Method from Gray-Level Histograms," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 9, no. 1, pp. 62-66, 1979.
- [38] A. K. Jain, "Data clustering: 50 years beyond K-means," *Pattern recognition letters*, vol. 31, no. 8, pp. 651-666, 2010.
- [39] K. R. Zalik, "An efficient k'-means clustering algorithm," *Pattern Recognition Letters*, vol. 29, no. 9, pp. 1385-1391, 2008.
- [40] H. Kato, and T. Harada, "Image reconstruction from bag-of-visual-words," pp. 955-962.
- [41] A. Bosch, X. Muñoz, and R. Martí, "Which is the best way to organize/classify images by content?," *Image and vision computing*, vol. 25, no. 6, pp. 778-791, 2007.
- [42] S. Xu, T. Fang, D. Li, and S. Wang, "Object classification of aerial images with bag-of-visual words," *IEEE Geoscience and Remote Sensing Letters*, vol. 7, no. 2, pp. 366-370, 2010.
- [43] M. Paul, R. K. Karsh, and F. A. Talukdar, "Image hashing based on shape context and speeded up robust features (SURF)," pp. 464-468.
- [44] F. Baig, Z. Mehmood, M. Rashid, M. A. Javid, A. Rehman, T. Saba, and A. Adnan, "Boosting the performance of the BoVW model using SURF-CoHOG-based sparse features with relevance feedback for CBIR," *Iranian Journal of Science and Technology, Transactions of Electrical Engineering*, vol. 44, no. 1, pp. 99-118, 2020.
- [45] M. Ghalan, and R. K. Aggarwal, "Multifold Classification for Human Action Recognition," pp. 1-6.
- [46] T. Sunitha, and T. Sivarani, "Novel content based medical image retrieval based on BoVW classification method," *Biomedical Signal Processing and Control*, vol. 77, pp. 103678, 2022.

- [47] E. R. Vimina, and K. P. Jacob, "Feature fusion method using BoVW framework for enhancing image retrieval," IET Image Processing, vol. 13, no. 11, pp. 1979-1985, 2019.
- [48] P. Panchal, S. Panchal, and S. Shah, "A comparison of SIFT and SURF," International Journal of Innovative Research in Computer and Communication Engineering, vol. 1, no. 2, pp. 323-327, 2013.
- [49] N. Y. Khan, B. McCane, and G. Wyvill, "SIFT and SURF Performance Evaluation against Various Image Deformations on Benchmark Dataset," pp. 501-506.
- [50] E. Karami, S. Prasad, and M. Shehata, "Image matching using SIFT, SURF, BRIEF and ORB: performance comparison for distorted images," arXiv preprint arXiv:1710.02726, 2017.
- [51] S. Routray, A. K. Ray, and C. Mishra, "Analysis of various image feature extraction methods against noisy image: SIFT, SURF and HOG," pp. 1-5.
- [52] L. Cao, J. Ling, and X. Xiao, "Study on the Influence of Image Noise on Monocular Feature-Based Visual SLAM Based on FFDNet," Sensors, vol. 20, no. 17, pp. 4922, 2020.
- [53] D. M. Farid, L. Zhang, C. M. Rahman, M. A. Hossain, and R. Strachan, "Hybrid decision tree and naïve Bayes classifiers for multi-class classification tasks," Expert Systems with Applications, vol. 41, no. 4, pp. 1937-1946, 2014.
- [54] N. Patel, and S. Upadhyay, "Study of various decision tree pruning methods with their empirical comparison in WEKA," International journal of computer applications, vol. 60, no. 12, 2012.
- [55] A. Priyama, R. G. Abhijeeta, A. Ratheeb, and S. Srivastavab, "Comparative analysis of decision tree classification algorithms," International Journal of Current Engineering and Technology, vol. 3, no. 2, pp. 334-337, 2013.
- [56] N. Bhatia, "Survey of nearest neighbor techniques," arXiv preprint arXiv:1007.0085, 2010.
- [57] T. Cover, and P. Hart, "Nearest neighbor pattern classification," IEEE transactions on information theory, vol. 13, no. 1, pp. 21-27, 1967.
- [58] J. Kim, B.-S. Kim, and S. Savarese, "Comparing image classification methods: K-nearest-neighbor and support-vector-machines," pp. 133-138.
- [59] A. Ben-Hur, and J. Weston, "A user's guide to support vector machines," Data mining techniques for the life sciences, pp. 223-239: Springer, 2010.
- [60] N. Reljin, and D. Pokrajac, "Classification of performers using support vector machines," pp. 165-169.
- [61] X. Yang, R. Zhang, Z. Zhai, Y. Pang, and Z. Jin, "Machine learning for cultivar classification of apricots (*Prunus armeniaca* L.) based on shape features," Scientia Horticulturae, vol. 256, pp. 108524, 2019.
- [62] J. Abellán, and J. G. Castellano, "Improving the Naive Bayes Classifier via a Quick Variable Selection Method Using Maximum of Entropy," Entropy, vol. 19, no. 6, pp. 247, 2017.
- [63] M. Ahmed, M. Shahjahan, M. Rana, M. Mollah, and N. Haque, "Robustification of Naïve Bayes Classifier and Its Application for Microarray Gene Expression Data Analysis," BioMed research international, vol. 2017, 2017.
- [64] S. Mukherjee, and N. Sharma, "Intrusion detection using naive Bayes classifier with feature reduction," Procedia Technology, vol. 4, pp. 119-128, 2012.
- [65] Y. Ozaki, M. Yano, and M. Onishi, "Effective hyperparameter optimization using Nelder-Mead method in deep learning," IPSJ Transactions on Computer Vision and Applications, vol. 9, no. 1, pp. 20, 2017.
- [66] M. F. Aslan, A. Durdu, A. Yusefi, K. Sabanci, and C. Sungur, "A tutorial: Mobile robotics, SLAM, bayesian filter, keyframe bundle adjustment and ROS applications," Robot Operating System (ROS), pp. 227-269, 2021.
- [67] M. F. Aslan, K. Sabanci, A. Durdu, and M. F. Unlarsen, "COVID-19 diagnosis using state-of-the-art CNN architecture features and Bayesian Optimization," Computers in Biology and Medicine, vol. 142, pp. 105244, 2022/03/01/, 2022.
- [68] A. Mikołajczyk, and M. Grochowski, "Data augmentation for improving deep learning in image classification problem," pp. 117-122.
- [69] C. Shorten, and T. M. Khoshgoftaar, "A survey on image data augmentation for deep learning," Journal of Big Data, vol. 6, no. 1, pp. 1-48, 2019.
- [70] J.-E. Kim, N.-E. Nam, J.-S. Shim, Y.-H. Jung, B.-H. Cho, and J. J. Hwang, "Transfer Learning via Deep Neural Networks for Implant

Fixture System Classification Using Periapical Radiographs," Journal of Clinical Medicine, vol. 9, no. 4, pp. 1117, 2020.

- [71] T. Hinz, N. Navarro-Guerrero, S. Magg, and S. Wermter, "Speeding up the Hyperparameter Optimization of Deep Convolutional Neural Networks," International Journal of Computational Intelligence and Applications, pp. 1850008, 2018.

BIOGRAPHY



MUHAMMET FATİH ASLAN received his B.Sc. and master's degrees from Selcuk University (SU), Konya, Turkey in 2016 and 2018, respectively. He completed his Ph.D. in Konya Technical University (KTUN) in 2022, in the department of electrical and electronics engineering. He has been working as a researcher at Karamanoglu Mehmetbey University (KMU) since 2017. His research interests include mobile robotics, real-time autonomous applications, image processing, machine learning, computer vision and object tracking.

Precoded Universal MIMO Superposition Transmission for Achieving Optimal Coverage and High Throughput in 6G and Beyond Networks

Sadiq Iqbal*, Jehad M. HAMAMREH

Abstract—A wireless network can utilize its entire resources to serve users without any allocation and provide upgraded performances such as enhanced throughput and high reliability without interference among users. However, previous wireless communication technologies could not bring the full utilization of network resources and overcome the challenge of forgoing scheduling among users. Therefore, to address these requirements, we propose and develop a novel design coined as Precoded Universal MIMO Superposition Transmission (PU-MIMO-ST) that can also be applicable for the most challenging and worst case interference scenario where the number of antenna points (APs) and the number of user equipment (UEs) is equal. In the considered system model, the APs are linked to a central processing unit (CPU) via backhaul and the UEs receive co-operation from the network resulting in achieving the optimal usage of network resources. The results obtained from computer simulations proof and verify the effectiveness of the proposed design called PU-MIMO-ST compared with other competitive works in terms of reliability, throughput, reduced complexity on the reception side, as well as power conservation.

Index Terms—Multi-user MIMO, massive MIMO, network MIMO, distributed MIMO, distributed antenna system, cell-free massive MIMO, pCell, 6G and Beyond.

I. INTRODUCTION

TO fulfill the communication needs of humans as well as intelligent machines, 6G and beyond networks will be the fundamental components for all parts of life, society, and industry. These networks will pave the way towards realizing the technological aspirations that include holographic telepresence, e-Health, ubiquitous connectivity in smart environments, massive robotics, three-dimensional massive unmanned mobility, augmented reality, virtual reality, and internet of everything [1]. To provide these applications with an effective and efficient wireless communication network than ever with an unparalleled increase in the data rates, massive connectivity, high-reliability, and low latency, 6G and beyond networks will implement promising technologies with different paradigms relevant to massive multiple input multiple output (MMIMO) networks such as network MIMO

(netMIMO), distributed MIMO (D-MIMO), multi-cell MIMO (MC-MIMO), distributed input distributed output (DIDO) and joint transmission coordinated multi-point (JT-CoMP) [1]. The conventional manifestation of the MMIMO system is the collocated MMIMO (C-MMIMO) in which a minority of users are served through a considerable majority of antennas via beamforming. The antennas are located at the base stations (BSs) and share the same temporal-frequency resources [2].

C-MMIMO systems can provide significantly high data rates, high link reliability, coverage, and energy efficiency [3], but suffer from performance degradation for cell-edge users located at cell edges due to the lower channel gain from the serving cell as well as the inter-cell interference. However, unlike the C-MMIMO system, the MMIMO operation can be employed in a distributed mode by arranging the majority of single antenna access points with a central processing unit (CPU). The CPU and access points are linked through a backhaul [4] and the setup is identical to a distributed antenna system (DAS) [5]. Distributed MMIMO (D-MMIMO) systems provide superior data rates than C-MMIMO due to the high diversity gain to each user and good quality-of-service (QoS) to cell-edge users as an outcome of having more APs on the edge users proximity.

A. Relevant D-MMIMO technologies

This subsection is dedicated to briefly overview and talk about the D-MMIMO related technologies like netMIMO, D-MIMO, MC-MIMO, DIDO, and JT-CoMP.

netMIMO: netMIMO can mitigate inter-cell interference (ICI) by synchronizing the access points and sharing data packets. The access points can transmit concurrently to increase the downlink range. A reference signal is utilized to synchronize the access points at the level of the carrier signal, and data is shared among the access points through a wired backhaul [6], [7], [8], [9], [10].

D-MIMO: A generalized DAS, also known as distributed MIMO (D-MIMO), contains the features of point-to-point MIMO and DAS while also achieving enhanced spatial diversity. Besides C-MIMO, the D-MIMO system suffers from path loss due to different distances from discretely scattered antennas, resulting in increasing difficulty of spectral efficiency analysis [11], [12].

MC-MIMO: Wyner presented a multi-cell system with inter-cell cooperation [13], where BSs with multiple antennas can form wireless networks around each other and each cell

Sadiq Iqbal* is with the WISLAB, Department of Electrical Computer Engineering, Antalya Bilim University, Antalya, 07190, Turkey (e-mail: sadiq.zahid73@gmail.com).

* Corresponding Author

Jehad M. HAMAMREH is with the WISLAB, Department of Electrical Electronics Engineering, Antalya Bilim University, Antalya, 07190, Turkey (e-mail: jehad.hamamreh@gmail.com).

Manuscript received Sep 20, 2022; accepted Dec 14, 2022.
DOI: 10.17694/bajece.1177534

All the codes used can be found at researcherstore.com

contains within its range a single-antenna user equipment (UE).

DIDO: DIDO network consists of access points and remote radio units (RRHs), which are utilized in human-type and machine-type communications (MTC) systems. The RRHs are relatively much nearer to UEs [14].

JT-CoMP: JT-CoMP consists of CoMP clusters that have distinct coordinating nodes that broadcast data to a UE consistently or non consistently, and phase information is utilized so that the network can remove inter-user interference [15], [16], [17], [18].

B. CF-MMIMO technologies

Recently, an updated practical version of the D-MMIMO system has emerged and named as cell-free MMIMO (CF-MMIMO) system. This system consists of a CPU connected to all the access points, operating them as an MMIMO network with no cell boundaries. By employing spatial multiplexing, the users are served over the same resources [1]. Even compared with the small-cell (SC) system, the CF-MMIMO system outperforms the fully distributed SC system in terms of the 95% likely per user throughput [19]. In the literature so far, there are many works done on the CF-MMIMO system such as, in [20], authors derive a closed-form expression, and through this expression, an optimal max-min power control scheme can be obtained which provides equivalent quality of service to all users.

Authors in [21] propose a new configuration for scalable CF-MMIMO systems by exploiting the dynamic cooperation cluster concept of netMIMO to scale joint initial access, allotment of pilots with the formation of clusters while having a duality of the uplink and downlink. Concerning the precoding and power optimization in CF-MMIMO, the authors in [22] propose a near-optimal power control algorithm that is considerably simpler than exact max-min power control. Allotment of power method discussed in [23] can increase the energy efficiency, and the authors also developed an access point selection method so that users can select access point subsets that reduce the backhaul link power consumption [23]. In [24], authors present an analysis of CF-MMIMO under distinct degrees of cooperation among the access points and the uplink spectral efficiencies of four different cell-free implementations, with spatially correlated fading and arbitrary linear processing.

Scalability aspects of CF-MMIMO have been investigated in [25]. In this study, the authors proposed a solution related to data processing, network topology, and power control. This framework achieves full scalability at the cost of a meager performance loss compared to the standard form of CF-MMIMO. In [26], authors assume how much downlink pilot can improve the performance of CF-MMIMO and determine that by exploiting downlink pilots in less dense networks, the enhancement in performance can be obtained while staying moderate in a high-density area.

Authors in [27] propose a beyond cell-free MIMO (CF-MMIMO) approach to surround the BSs and users with reconfigurable intelligent surfaces (RISs). In this less costly way,

reconfigurable reflections can produce spreading conditions, thus enhancing CF-MMIMO communication. The study [28] takes a new look at CF-MMIMO via the lens of the dynamic cooperation cluster framework from netMIMO while providing distributed algorithms for merging, initial access, precoding, cluster formation, and assignment of pilots to make them expandable for multiple users.

Another important commercialized application relevant to CF-MMIMO but better in terms of spectral efficiency and presented under the umbrella of the Artemis network is pCell [29] pCell is all about procuring SE multiple times of any other technology by distributing both the transmitters and receivers. The transmitters are scattered while the receivers are packed in a small space. Taking advantage of the interference, coordination between access points and UEs, and equipped with peak spatially multiplexed yields acquired through numerous transmissions. SE gains are reaped without any concept of existing cells [29]. Authors in [29] also consider a unique geometrical spreading channel model that is customized as a merger of plane waves which is different from the one utilized in [19].

So far, we have concisely discussed MMIMO with its relevant distributed technologies, and then we talked about the CF-MMIMO with its technical literature review and the pCell. Given below are the demerits of D-MMIMO and CF-MMIMO that highlight their drawbacks, and then we bring forth our proposed model's advantages that remedy these drawbacks.

C. Shortcomings of D-MMIMO

- 1) D-MMIMO utilize non-cell free approach that makes it harder to achieve total coverage and geographical area flexibility.
- 2) To eliminate ICI, synchronization is needed among access points which increases network's complexity.
- 3) Due to multi-cell scenario, users at the cell edges experience severe degradation in performances and throughput.
- 4) Due to fixed inter-cell cooperation, the resources of the network are divided among users.

D. Shortcomings of CF-MMIMO

- 1) There are no cells present in CF-MMIMO with access points scattered around the perceivable serving area. The number of access points, however, are much larger than the users, leading to uneven sharing of network resources among users.
- 2) Needs scheduling to accommodate users utilizing the resources.
- 3) Super positioning of user data not implemented in CF-MMIMO.

As can be noticed, the aforementioned state-of-the-art technologies do not provide capacity usage of the system with unprecedented spectral efficiency, highest reliability, and interference-free communication to each UE at the same time. These challenges need to be addressed for the future 6G and beyond networks. To meet these requirements and fulfill the shortcomings of the D-MMIMO and CF-MMIMO, in this work, we propose a new novel communication technique.

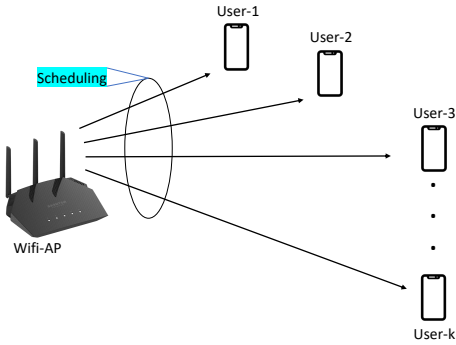


Fig. 1. Representation of a conventional WiFi access points serving multiple users.

E. Contributions of the proposed design

The contributions provided by the proposed design in this work are summarized below:

- 1) The proposed work can directly serve each user with all the resources of a cell in a network which includes several antennas, time slots, spectrum bandwidth, etc., without causing any interference at all.
- 2) The number of APs is made equal to the number of UEs, which results in dedicating the capacity of the network to every UE simultaneously.
- 3) The method is enabled by superimposing the users' data in a very unique and special way to eliminate multiuser interference.
- 4) Specially designed diagonal matrices are utilized along with novel precoding and auxiliary signals that are a function of both user data and the channel effects.

F. Differences & impact of the proposed method

To explain the differences between the proposed design and the state-of-the-art techniques, we consider an example of a WiFi access point as shown in Fig. 1. The access point is installed in a room to initially serve just a single user, who is assumed to be served with a data rate of 50 Mbps for instance. However, when a new user gets connected to the same WiFi access point, the system now communicates with two users and splits the resources and thus the data rate in half so that both users can get 25 Mbps each. The WiFi will split the data according to the number of users associated with it. The WiFi does the splitting of resources in the time domain due to the WiFi spectrum being overly crowded and full of interference.

Scheduling is the same concerning 4G and 5G technologies, but compared with WiFi is much better because cellular systems utilize time and frequency to assign time slots and carriers to a user based on their conditions and service requirements such as text, voice, gaming, video streaming, etc. In scheduling, the process of adjusting, adapting, fixing, redistributing, and reallocating the resources continues to happen as long as there are users trying to get access simultaneously. Therefore, in wireless systems, gaining full optimal capacity is the best-case scenario, which results in allocating the full bandwidth of the access point, BS, or the

tower to every user in the whole network without sharing and with no interference at the user location. One of the main contributions of the proposed design is exactly that as the users number increases, they get the full bandwidth with no splitting or sharing of resources with other users. This makes the proposed design clearly different from CF-MMIMO [19] and pCell [29].

The rest of the paper is structured as follows. Section II expounds the considered system model framework, the channel representation, the proposed theorem that incorporates the introduced superposition concept, and derives the closed-form expressions of the superimposed auxiliary signals and precoding matrices. Section III highlights the performance inspection and measures of the proposed design. Section IV illustrates the simulation results and explains them. Finally, Section V presents the conclusion of the paper¹.

II. SYSTEM LAYOUT

The layout of the proposed system is as shown in Fig. 2, where we consider the number of UEs denoted by K equal to the number of APs signified by M . The APs are connected to a CPU via a backhaul, and both users and access points are placed irregularly in a geographical cell-reprieved environment as shown in Fig. 2, which depicts the general illustration of the proposed scheme comprising three different scenarios ranging from two UEs case to K UEs case and number of K UEs is equal to number of M APs.

A. Transmission

The APs simultaneously transmit to all UEs for achieving the spatial multiplexing property. The data of all UEs are precoded, superimposed on top of each other, and then the auxiliary signals are added to them. After this, the Inverse Fast Fourier Transform (IFFT) is applied to the signals to be transmitted.

B. Channel composition

The complex channel coefficient is styled as the superimposition of plane waves across AP antenna $m=1,...,M$ and UE antenna $k=1,...,K$

$$h_{km} = \sum_{p \in S_{km}} d_p e^{-i\omega \hat{v}_p \cdot r_k} \quad (1)$$

In Eq. (1), every path is included in the set S_{km} , such as propagation and cluster scattering through AP antenna m to UE antenna k . Wavenumber $\omega = 2\pi/\lambda$, λ denotes the wavelength. As shown in Fig. 4, the placement vector of UE k is r_k relative to an origin O , while the unit vector is \hat{v}_p which is in the direction of the incident path p which points out from the position of UE k . d_p on the other hand, is the complex coefficient with specifically designed path loss, shadowing, and phase terms that are unassociated to r_k for path p .

¹Notations: Vectors are expressed and denoted by bold-small letters, whereas matrices are expressed by bold-large letters. The transpose, conjugate transpose, and inverse are symbolized by $(\cdot)^T$, $(\cdot)^H$ and $(\cdot)^{-1}$, respectively. \mathbf{I} is the $N \times N$ identity matrix and \mathbf{O} is the $N \times N$ Zero matrix.

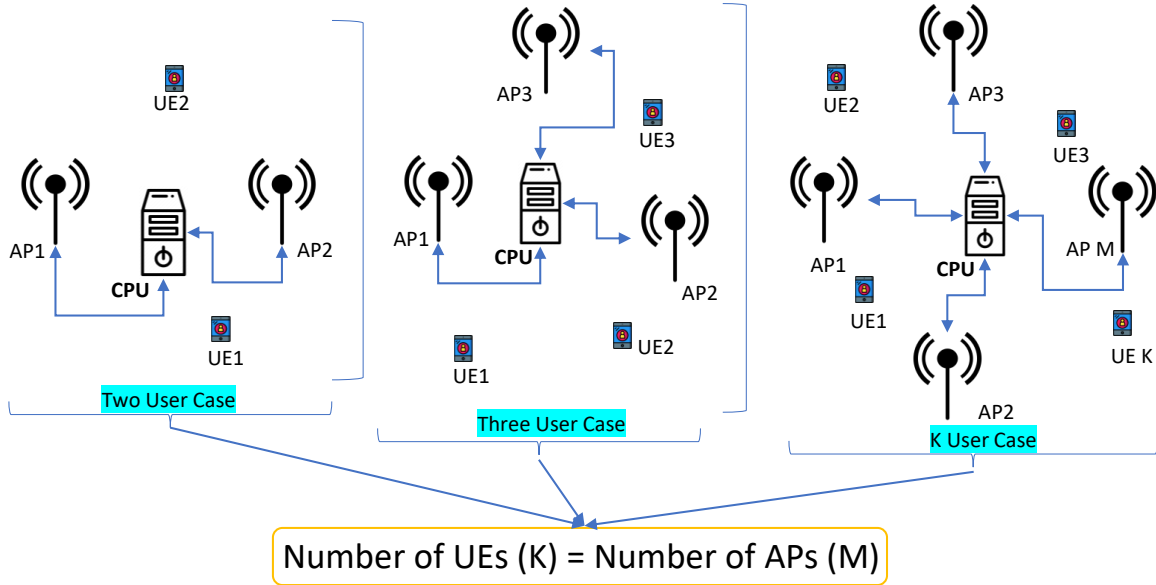


Fig. 2. General system model of the proposed method.

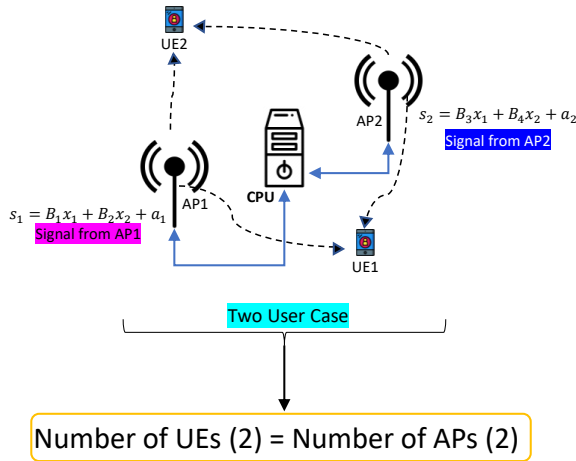
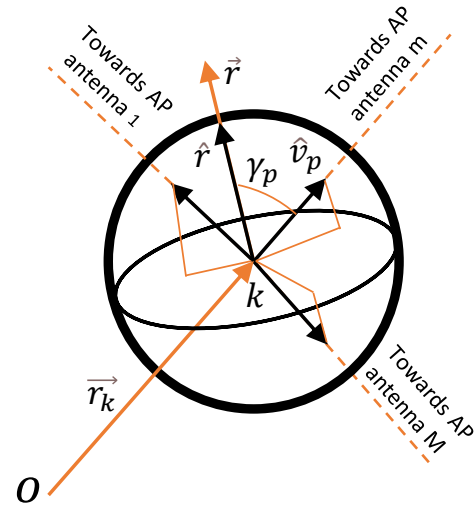


Fig. 3. System model of the proposed method with two access points and two users.

If we exclude d_p path loss from Eq. 1, we are left with the exponent component, which houses the small scale fading. The average SNR consists of the path loss effect that is received by the UE and the remaining channel components boil down to a Rayleigh fading phenomenon as in [30], [31], [32].

C. Reception

As illustrated in Fig. 3, each of the AP_m and UE_k has a single antenna. The channel matrix detailing the propagation between the transceiver antennas is denoted by H_{km} . The total number of modulated symbols in one OFDM block at each AP for each UE is N_f , thus the frequency response of every OFDM block for each UE is given by

Fig. 4. Design parameters of the channels with angle $\gamma_p \in [0, \pi]$ across directional path \hat{v}_p and direction of displacement \hat{r} .

$\mathbf{x}_k = [\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_{N_f-1}] \in C^{[N_f \times 1]}$. The auxiliary signal [33] is denoted by \mathbf{a}_q , $q = m$ and precoding matrix [34] indicated by \mathbf{B}_t , $t = 1^2, 2^2, \dots, k^2$ both of which are based on the wireless channel propagation properties. The three case studies will be discussed and the corresponding conditions are illustrated that can be used to find the values of precoders and auxiliary signals.

1) In case of UE_k with $k = 1, 2$: As we have presumed earlier $UE_k = AP_m$, AP_1 and AP_2 concurrently transmit the intended superimposed signals s_1 and s_2 to UE_1 and UE_2 . The superimposed signals are as follows:

$$s_1 = \mathbf{B}_1 \mathbf{x}_1 + \mathbf{B}_2 \mathbf{x}_2 + \mathbf{a}_1, \quad (2)$$

$$\mathbf{s}_2 = \mathbf{B}_3\mathbf{x}_1 + \mathbf{B}_4\mathbf{x}_2 + \mathbf{a}_2, \quad (3)$$

where $\mathbf{B}_1, \mathbf{B}_2, \mathbf{B}_3$ and \mathbf{B}_4 are precoding matrices, while \mathbf{a}_1 and \mathbf{a}_2 are the superimposed auxiliary signals as shown in (2) and (3).

Signals received at \mathbf{UE}_1 :

$$\mathbf{y}_{11} = \mathbf{H}_{11}\mathbf{s}_1, \quad (4)$$

$$\mathbf{y}_{12} = \mathbf{H}_{12}\mathbf{s}_2, \quad (5)$$

$$\hat{\mathbf{y}}_{r1} = \mathbf{y}_{11} + \mathbf{y}_{12} + \mathbf{z}_1, \quad (6)$$

$$\hat{\mathbf{y}}_{r1} = \mathbf{H}_{11}\mathbf{s}_1 + \mathbf{H}_{12}\mathbf{s}_2 + \mathbf{z}_1, \quad (7)$$

$$\begin{aligned} \hat{\mathbf{y}}_{r1} &= \mathbf{H}_{11}(\mathbf{B}_1\mathbf{x}_1 + \mathbf{B}_2\mathbf{x}_2 + \mathbf{a}_1) \\ &+ \mathbf{H}_{12}(\mathbf{B}_3\mathbf{x}_1 + \mathbf{B}_4\mathbf{x}_2 + \mathbf{a}_2) + \mathbf{z}_1, \end{aligned} \quad (8)$$

$$\begin{aligned} \hat{\mathbf{y}}_{r1} &= (\mathbf{B}_1\mathbf{H}_{11} + \mathbf{B}_3\mathbf{H}_{12})\mathbf{x}_1 + (\mathbf{B}_2\mathbf{H}_{11} + \mathbf{B}_4\mathbf{H}_{12})\mathbf{x}_2 \\ &+ \mathbf{H}_{11}\mathbf{a}_1 + \mathbf{H}_{12}\mathbf{a}_2 + \mathbf{z}_1. \end{aligned} \quad (9)$$

where \mathbf{y}_{11} and \mathbf{y}_{12} are the intended signals for \mathbf{UE}_1 , $\hat{\mathbf{y}}_{r1}$ is the combined received signal at \mathbf{UE}_1 . \mathbf{H}_{11} and \mathbf{H}_{12} are the channel impulse responses between \mathbf{UE}_1 and both \mathbf{AP}_1 and \mathbf{AP}_2 , while \mathbf{z}_1 is the additive white Gaussian noise present at \mathbf{UE}_1 antenna.

Signals received at \mathbf{UE}_2 :

$$\mathbf{y}_{21} = \mathbf{H}_{21}\mathbf{s}_1, \quad (10)$$

$$\mathbf{y}_{22} = \mathbf{H}_{22}\mathbf{s}_2, \quad (11)$$

$$\hat{\mathbf{y}}_{r2} = \mathbf{y}_{21} + \mathbf{y}_{22} + \mathbf{z}_2, \quad (12)$$

$$\hat{\mathbf{y}}_{r2} = \mathbf{H}_{21}\mathbf{s}_1 + \mathbf{H}_{22}\mathbf{s}_2 + \mathbf{z}_2, \quad (13)$$

$$\begin{aligned} \hat{\mathbf{y}}_{r2} &= \mathbf{H}_{21}(\mathbf{B}_1\mathbf{x}_1 + \mathbf{B}_2\mathbf{x}_2 + \mathbf{a}_1) \\ &+ \mathbf{H}_{22}(\mathbf{B}_3\mathbf{x}_1 + \mathbf{B}_4\mathbf{x}_2 + \mathbf{a}_2) + \mathbf{z}_2, \end{aligned} \quad (14)$$

$$\begin{aligned} \hat{\mathbf{y}}_{r2} &= (\mathbf{B}_1\mathbf{H}_{21} + \mathbf{B}_3\mathbf{H}_{22})\mathbf{x}_1 + (\mathbf{B}_2\mathbf{H}_{21} + \mathbf{B}_4\mathbf{H}_{22})\mathbf{x}_2 \\ &+ \mathbf{H}_{21}\mathbf{a}_1 + \mathbf{H}_{22}\mathbf{a}_2 + \mathbf{z}_2. \end{aligned} \quad (15)$$

where \mathbf{y}_{21} and \mathbf{y}_{22} are the intended signals for \mathbf{UE}_2 , $\hat{\mathbf{y}}_{r2}$ is the combined received signal at \mathbf{UE}_2 . \mathbf{H}_{21} and \mathbf{H}_{22} are the channel impulse responses between \mathbf{UE}_2 and both \mathbf{AP}_1 and \mathbf{AP}_2 , while \mathbf{z}_2 is the additive white Gaussian noise available at \mathbf{UE}_2 antenna.

2) In case of \mathbf{UE}_k and \mathbf{AP}_m : The generalized closed form expressions of the transmitted and received signals are as follows.

$$\mathbf{s}_m = \sum_{k=1}^K \mathbf{B}_t\mathbf{x}_k + \mathbf{a}_q, \quad (16)$$

$$\hat{\mathbf{y}}_{rk} = \sum_{m=1}^M \mathbf{H}_{km}\mathbf{s}_m + \mathbf{z}_k. \quad (17)$$

Eqs. (16) and (17) represent the transmitted \mathbf{s}_m and received $\hat{\mathbf{y}}_{rk}$ signals for the k_{th} user. In Eq. (16), the \mathbf{B}_t is the designed precoder, \mathbf{x}_k denotes the user data of the k_{th} user, and \mathbf{a}_q is the auxiliary signal. Eq. (17) includes \mathbf{H}_{km} , \mathbf{s}_m and \mathbf{z}_k that are the channel impulse response between the k_{th} user and m_{th} \mathbf{AP} , m_{th} signal and the additive white Gaussian noise at the k_{th} \mathbf{UE} .

3) Designing the conditional system of equations for precoders and auxiliary signals: For a user (\mathbf{UE}_k) to receive its data, the interference from other \mathbf{UE}_s and the channel effects on this user must be removed. This outcome can be obtained by the system of equations given below.

In the case of two users, the precoding conditions for each user are to equate the required data expression to the identity matrix and the unwanted terms to the null matrix. From Eq. (9), the term associated with \mathbf{x}_1 is equated to the identity matrix and the term linked to \mathbf{x}_2 is equated to a null matrix such as:

$$\mathbf{B}_1\mathbf{H}_{11} + \mathbf{B}_3\mathbf{H}_{12} = \mathbf{I}, \quad (18)$$

$$\mathbf{B}_2\mathbf{H}_{11} + \mathbf{B}_4\mathbf{H}_{12} = \mathbf{O}, \quad (19)$$

From Eq. (15), the term multiplying \mathbf{x}_1 is equated to the null matrix and the expression multiplying \mathbf{x}_2 is equated to an identity matrix like:

$$\mathbf{B}_1\mathbf{H}_{21} + \mathbf{B}_3\mathbf{H}_{22} = \mathbf{O}, \quad (20)$$

$$\mathbf{B}_2\mathbf{H}_{21} + \mathbf{B}_4\mathbf{H}_{22} = \mathbf{I}, \quad (21)$$

Therefore, by following Eqs. (18), (19), (20), and (21), the conditions for deriving the precoders for two users with $k = 1, 2$ can be written. Eqs. (18) and (20) can be combined to form a system of equations (22) and by solving this system the values of precoders \mathbf{B}_1 and \mathbf{B}_3 can be determined.

$$\begin{cases} \mathbf{B}_1\mathbf{H}_{11} + \mathbf{B}_3\mathbf{H}_{12} = \mathbf{I}, \\ \mathbf{B}_1\mathbf{H}_{21} + \mathbf{B}_3\mathbf{H}_{22} = \mathbf{O}, \end{cases} \quad (22)$$

Similar to Eqs. (18) and (20), by combining Eqs. (21) and (19) to form a system of equations (23), the values of precoders \mathbf{B}_2 and \mathbf{B}_4 can be calculated.

$$\begin{cases} \mathbf{B}_2\mathbf{H}_{21} + \mathbf{B}_4\mathbf{H}_{22} = \mathbf{I}, \\ \mathbf{B}_2\mathbf{H}_{11} + \mathbf{B}_4\mathbf{H}_{12} = \mathbf{O}. \end{cases} \quad (23)$$

The above precoder conditions illustrate systems of equations that represent the pattern of increase in equations in the system and the increase of the system as a whole. For $k = 1, 2$, there are two systems of equations with a pair of equations

in each system, this leads to a generalization of the system of equations for k_{th} user such as:

$$n = 2, 3, \dots, N \rightarrow \begin{cases} \sum_{m=1}^M \mathbf{B}_t \mathbf{H}_{km} = \mathbf{I}, \\ \sum_{m=1}^M \mathbf{B}_t \mathbf{H}_{km} = \mathbf{O}. \end{cases} \quad (24)$$

where $n = 2, 3, \dots, N$ is the increase in number of equations inside the system of equations starting from two and $N \equiv k_{th} \text{ user}$.

Now the conditional system of equations for auxiliary signals will be derived for two users with $k = 1, 2$. From Eq. (9), for the first user the first term is the required one, the second, third, and fourth terms are equated to the null matrix, like:

$$(\mathbf{B}_2 \mathbf{H}_{11} + \mathbf{B}_4 \mathbf{H}_{12}) \mathbf{x}_2 + \mathbf{H}_{11} \mathbf{a}_1 + \mathbf{H}_{12} \mathbf{a}_2 = \mathbf{O}, \quad (25)$$

From Eq. (15), for the second user, the second term is the desired term, thus the first, third, and fourth terms are equated to the null matrix as shown below:

$$(\mathbf{B}_1 \mathbf{H}_{21} + \mathbf{B}_3 \mathbf{H}_{22}) \mathbf{x}_1 + \mathbf{H}_{21} \mathbf{a}_1 + \mathbf{H}_{22} \mathbf{a}_2 = \mathbf{O}. \quad (26)$$

Now for calculating the auxiliary signals the conditional system of equations can be derived by combining the Eqs. (25) and (26) as:

$$\begin{cases} (\mathbf{B}_2 \mathbf{H}_{11} + \mathbf{B}_4 \mathbf{H}_{12}) \mathbf{x}_2 + \mathbf{H}_{11} \mathbf{a}_1 + \mathbf{H}_{12} \mathbf{a}_2 = \mathbf{O}. \\ (\mathbf{B}_1 \mathbf{H}_{21} + \mathbf{B}_3 \mathbf{H}_{22}) \mathbf{x}_1 + \mathbf{H}_{21} \mathbf{a}_1 + \mathbf{H}_{22} \mathbf{a}_2 = \mathbf{O}. \end{cases} \quad (27)$$

Eq. (27) represents the conditional system of equations for finding the values of auxiliary signals for two users case. By following these assumptions, we can formalize the general form of a system of equations as follows:

$$n \begin{cases} \sum_{c=n-k}^K ((\sum_{m=1}^M \mathbf{B}_t \mathbf{H}_{km}) \mathbf{x}_k) + \sum_{m=1}^M \mathbf{H}_{km} \mathbf{a}_k = \mathbf{O}, \\ \sum_{c=n-k}^K ((\sum_{m=1}^M \mathbf{B}_t \mathbf{H}_{km}) \mathbf{x}_k) + \sum_{m=1}^M \mathbf{H}_{km} \mathbf{a}_k = \mathbf{O}. \end{cases} \quad (28)$$

The null matrix \mathbf{O} is equated to every equation in the system of equations to fulfill the conditions required for the calculation of auxiliary signals.

Designing precoders for the proposed method:

From Eq. (22) the values of precoders \mathbf{B}_1 and \mathbf{B}_3 can be derived by using the mathematical substitution method.

$$\mathbf{B}_1 = -\mathbf{H}_{22}(\mathbf{H}_{12} \mathbf{H}_{21} - \mathbf{H}_{11} \mathbf{H}_{22})^{-1}, \quad (29)$$

$$\mathbf{B}_3 = \mathbf{H}_{21}(\mathbf{H}_{12} \mathbf{H}_{21} - \mathbf{H}_{11} \mathbf{H}_{22})^{-1}, \quad (30)$$

For precoders \mathbf{B}_2 and \mathbf{B}_4 , the system of equations in (23) can be solved to get the precoder values such as:

$$\mathbf{B}_2 = -\mathbf{H}_{12}(\mathbf{H}_{11} \mathbf{H}_{22} - \mathbf{H}_{12} \mathbf{H}_{21})^{-1}, \quad (31)$$

$$\mathbf{B}_4 = \mathbf{H}_{11}(\mathbf{H}_{11} \mathbf{H}_{22} - \mathbf{H}_{12} \mathbf{H}_{21})^{-1}. \quad (32)$$

Designing auxiliary signals for the proposed method:

For finding the values of auxiliary signals in the case of two UEs, the conditional system of equations such as (27) will

be concurrently solved by substitution method to obtain the required solutions for \mathbf{a}_1 and \mathbf{a}_2 as:

$$\mathbf{a}_2 = (\mathbf{H}_{11}(\mathbf{B}_1 \mathbf{H}_{21} + \mathbf{B}_3 \mathbf{H}_{22}) \mathbf{x}_1 - \mathbf{H}_{21}(\mathbf{B}_2 \mathbf{H}_{11} + \mathbf{B}_4 \mathbf{H}_{12}) \mathbf{x}_2)(\mathbf{H}_{12} \mathbf{H}_{21} - \mathbf{H}_{11} \mathbf{H}_{22})^{-1}, \quad (33)$$

$$\mathbf{a}_1 = (-\mathbf{H}_{12}(\mathbf{B}_1 \mathbf{H}_{21} + \mathbf{B}_3 \mathbf{H}_{22}) \mathbf{x}_1 - \mathbf{H}_{22} \mathbf{a}_2)(\mathbf{H}_{21})^{-1}. \quad (34)$$

III. PERFORMANCE INSPECTION

This section is dedicated to the inspection of the performance analysis of the proposed technique. The proposed method's SNR is the function of the instantaneous power of the effective corresponding channel. Therefore, we are required to calculate the distributions associated with this quantity to find the distribution related to P_{γ_b} .

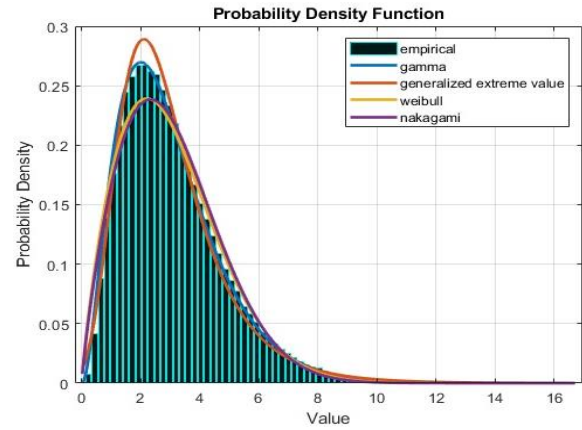


Fig. 5. Amplitude (Power) distribution of the effective fading channel of UE_1 .

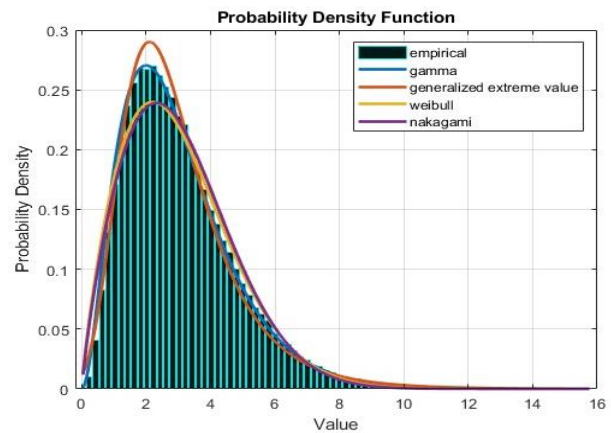


Fig. 6. Amplitude (Power) distribution of the effective fading channel of UE_2 .

A. Bit error rate (BER) analysis

To analyze the BER performance of the proposed scheme, since BPSK is used as a modulation in the scheme, the BER can be expressed in terms of [35] as

$$BER_{UE_k} = \frac{1}{2} \int_0^\infty \text{erfc}(\sqrt{\gamma_b}) P_{\gamma_b}(\gamma_b) d\gamma_b \quad (35)$$

where $\text{erfc}(\cdot)$ denotes the complementary error function, $P_{\gamma_b}(\gamma_b)$ is concerned with instantaneous SNR of the appropriate user as well as fading distribution and is termed as probability distribution function (PDF) as in [36].

$$P_{\gamma_b}(\gamma_b) = \frac{P_\alpha(\sqrt{\frac{\Omega\gamma_b}{\gamma_b}})}{2\sqrt{\frac{\gamma_b\gamma_b}{\Omega}}} \quad (36)$$

In Eq. (36), α denotes the effective channel of the concerned UE. For user 1 the effective channel is shown below:

$$\mathbf{B}_1\mathbf{H}_{11} + \mathbf{B}_3\mathbf{H}_{12}, \quad (37)$$

Similarly, for user 2, the effective is given below:

$$\mathbf{B}_2\mathbf{H}_{21} + \mathbf{B}_4\mathbf{H}_{22}. \quad (38)$$

In Eq. (36), Ω is the mean square of the channel fading amplitude $\Omega = \mathbb{E}\{\alpha^2\}$ while $\bar{\gamma}_b$ is the average SNR. Figs. 5 and 6 illustrate the distributions of effective fading channels of the two users, different distributions are fitted by using data fitting methods. From Figs. 5 and 6, Gamma distribution is the best fit due to the effective channels of the both users with shape and scale parameters a and b . The fading distribution can be described by the following, (D_α) .

$$(D_\alpha) \begin{cases} P_\alpha(\alpha) = \frac{1}{b^a\Gamma(a)}\alpha^{(a-1)}\exp(\frac{-\alpha}{b}) \\ (a_1, b_1) = (2.9849, 1.0069) \\ (a_2, b_2) = (2.9876, 1.0035) \end{cases} \quad (39)$$

Eq. (39) represents the expression for the fading profile of the effective channel of one user (denoted by shape and scale parameters a_1 and b_1 and a_2 and b_2), now the effective SNR PDF can be written as:

$$\begin{cases} P_{\gamma_b}(\gamma_b) = G\sqrt{\gamma_b}^{(a-2)}\exp(-(\frac{\sqrt{\Omega}}{b\sqrt{\gamma_b\gamma_b}})\gamma_b) \\ G = \frac{1}{2b^a\Gamma(a)}\sqrt{\frac{\Omega^a}{\gamma_b^a}} \end{cases} \quad (40)$$

where $\Gamma(\cdot)$ is the gamma function.

To calculate the BER including $\text{erfc}(\cdot)$ for the required user PDF profile, we substitute (40) in the BER $\text{erfc}(\cdot)$ expression (35), then we get the following BER integral form as

$$BER_{\gamma_b} = G\frac{1}{2}\int_0^\infty \text{erfc}(\sqrt{\gamma_b})\sqrt{\gamma_b}^{(a-2)} \exp(-(\frac{\sqrt{\Omega}}{b\sqrt{\gamma_b\gamma_b}})\gamma_b) d\gamma_b \quad (41)$$

The integral in Eq. (41) is of the same format presented in [37] and is given as:

$$\int_0^\infty \text{erfc}(\sqrt{x})\sqrt{x}\exp(-\beta x)dx = \frac{1}{2\sqrt{\pi}}\left(\frac{\arctan(\sqrt{\beta})}{(\beta)^{3/2}} - \frac{1}{2\beta(1+\beta)}\right), \beta > 0. \quad (42)$$

In Eq. (42) $\arctan(\cdot)$ denotes the inverse tangent function. To write the BER equation in (41) in the form of this integral by associating $a-2=1$ and appointing

$$\beta = -(\frac{\sqrt{\Omega}}{b\sqrt{\gamma_b\gamma_b}})\gamma_b. \quad (43)$$

It is possible to solve the concerning integration which leads to an approximate expression for the BER given as

$$BER_{\gamma_b} \approx \frac{G}{4\sqrt{\pi}}\left(\frac{\arctan(\sqrt{\beta})}{(\beta)^{3/2}} - \frac{1}{2\beta(1+\beta)}\right) \quad (44)$$

Under change of parameters Eq. (44) gives an approximate expression for BER of \mathbf{UE}_1 (and the same for \mathbf{UE}_2).

B. Evaluation of signal to interference plus noise ratio (SINR) for the proposed method

Due to the proposed system model, the interference caused by the data super-position and the channel effects experienced by the concerned user data are removed by the precoder matrices and auxiliary signals. The signals that are received at \mathbf{UE}_1 and \mathbf{UE}_2 expressed by the equations (9) and (15) after removing interference and elimination of channel effects can be further simplified as:

$$\mathbf{y}_{r1} = \mathbf{x}_1 + \mathbf{z}_1 \quad (45)$$

$$\mathbf{y}_{r2} = \mathbf{x}_2 + \mathbf{z}_2 \quad (46)$$

Eqs. (45) and (46) give the intended data symbols for the related \mathbf{UE}_s . Now, the following derivations will be related to the SINR for the proposed framework. The metrics imposing for the \mathbf{UE}_1 and \mathbf{UE}_2 's SINRs are as follows:

$$SINR_{\mathbf{UE}_k} = \frac{S_{p_m}}{I_{p_k} + \sigma_z^2} \quad (47)$$

where S_{p_m} , I_{p_k} and σ_z^2 denote the signal power, the interference power and the noise variance for the concerned \mathbf{UE} . SINR for \mathbf{UE}_1 and \mathbf{UE}_2 can be derived from the equations (45) and (46) and written as:

$$SINR_{\mathbf{UE}_1} = \frac{\|I_1\|^2}{\sigma_{z_1}^2} \quad (48)$$

$$SINR_{\mathbf{UE}_2} = \frac{\|I_2\|^2}{\sigma_{z_2}^2} \quad (49)$$

where I_1 and I_2 denoted the identity matrices.

IV. SIMULATED RESULTS

Simulation results showing the proposed method's performance in terms of BER, throughput error rate (TER), and peak to average power ratio (PAPR) are displayed and discussed in this section. The inspection of the system performance is based on the Rayleigh fading channel model as depicted in Table 1.

In Fig. 7, the BER performances of the proposed design, conventional MIMO systems, and proposed system utilizing auxiliary signals (AS) and precoding (PR) methods independently are presented. The BER performances $BER_1 - (\mathbf{AP} = 2)$ and $BER_2 - (\mathbf{AP} = 2)$ with \mathbf{AP} s equal to two are higher than MIMO (nTx=2 and nRx=2) with

TABLE I
PARAMETERS UTILIZED.

Type of Modulation	BPSK
IFFT / FFT size	64
Channel model	Rayleigh
Channel delay samples	[0 3 5 6 8]
Channel tap profile (dBm)	[0 -8 -17 -21 -25]
Guard interval length	16
Number of iterations	20000
Symbols per frame	1
Number of bits per second	1
Modulation order	2

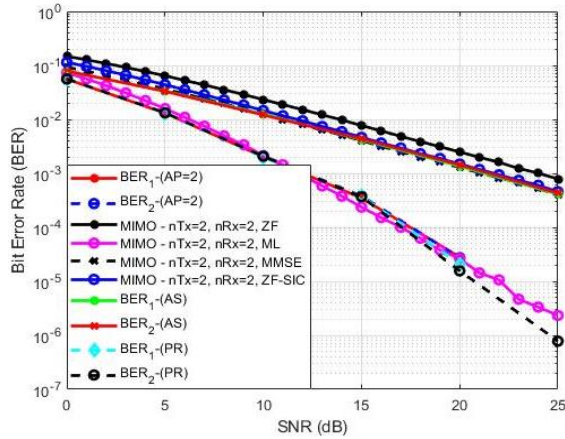


Fig. 7. Bit error rate: Showing the performances of proposed and conventional MIMO systems.

zero-forcing (ZF), maximum likelihood (ML), minimum mean square error (MMSE), and ZF with successive interference cancellation (ZF-SIC) methods. This comparison illustrates the effectiveness of the proposed design.

We can also see the comparison of the illustrations of the BER performances $BER_1 - AS$, $BER_2 - AS$ representing the system using AS superposition method and $BER_1 - PR$, $BER_2 - PR$ representing system utilizing PR method with conventional MIMO systems mentioned above. The AS method has the identical BER performance to ZF, MMSE, and ZF-SIC methods but is less than ML. However, due to its unique design, the PR technique has higher BER performance than all the conventional methods. Therefore, the proposed method's BER performance is the culmination of both AS and PR methods evolving into overall higher performance as depicted in Fig. 7.

Fig. 8 illustrates the TER performances of the proposed design $TER_1 - (AP = 2)$, $TER_2 - (AP = 2)$, total $TER - (AP = 2)$, and conventional MIMO ZF, ML, MMSE, and ZF-SIC systems. The proposed scheme has a higher TER than all the MIMO methods portrayed in Fig. 8. The total TER gain result is shown, which is double that of the individual TER_1 , TER_2 , and conventional MIMO TER performances generated by the super-positioned method of the proposed technique.

Moreover, Fig. 9 displays the PAPR gain of the proposed technique compared to the Conv. OFDM, system. The findings show that the proposed design can save power and is less complex than the conventional wireless systems.

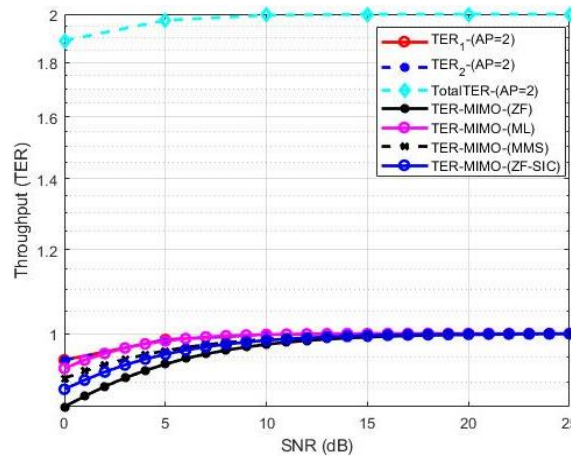


Fig. 8. Throughput error rate: Showing the TER performances for proposed and conventional MIMO systems.

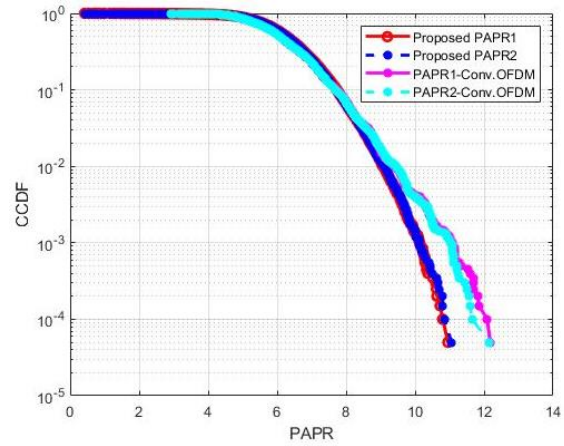


Fig. 9. Peak to Average Power Ratio: Illustrates the performances of proposed model ($PAPR_1$ and $PAPR_2$) and conventional OFDM system.

V. CONCLUSION

In this paper, we introduced and demonstrated a novel transmission scheme called PU-MIMO-ST that allows the optimal full utilization of network resources by each user, thus resulting in achieving high throughput gain thanks to the specially designed method of super-positioning the users data before transmission. The AS and PR matrices cancel the inter-antenna, inter-user interference, and channel effects before reaching the receiver, thus producing outcomes such as less complexity and processing at each UE. The obtained computer simulation results prove that the performance of the proposed system is better than the individual systems utilizing AS, PR, and conventional MIMO based systems. We also intend to extend the proposed technique to existing papers [36], [38], [39], [34], [40], [41], [42], [43], [44], [45], [46], [47], [48], [49], [50], [51], [52] that are related to signal superposition, auxiliary signals, and precoding methods.

REFERENCES

- [1] S. Elhoushy, M. Ibrahim, and W. Hamouda, "Cell-free massive MIMO: A survey," *IEEE Communications Surveys & Tutorials*, vol. 24, no. 1, 2021.

- [2] M. A. Albreem, M. Juntti, and S. Shahabuddin, "Massive MIMO detection techniques: A survey," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 4, pp. 3109–3132, 2019.
- [3] N. Fatema, G. Hua, Y. Xiang, D. Peng, and I. Natgunanathan, "Massive MIMO linear precoding: A survey," *IEEE systems journal*, vol. 12, no. 4, pp. 3920–3931, 2017.
- [4] J. Wu, Z. Zhang, Y. Hong, and Y. Wen, "Cloud radio access network (C-RAN): a primer," *IEEE network*, vol. 29, no. 1, pp. 35–41, 2015.
- [5] S. Elhoshy, M. Ibrahim, M. Ashour, T. Elshabrawy, H. Hammad, and M. M. Rizk, "A dimensioning framework for indoor DAS LTE networks," in *2016 International Conference on Selected Topics in Mobile & Wireless Networking (MoWNeT)*. IEEE, 2016, pp. 1–8.
- [6] X. Zhang, K. Sundaresan, M. A. Khojastepour, S. Rangarajan, and K. G. Shin, "NEMOx: Scalable network MIMO for wireless networks," in *Proceedings of the 19th annual international conference on Mobile computing & networking*, 2013, pp. 453–464.
- [7] D. Gesbert, S. Hanly, H. Huang, S. S. Shitz, O. Simeone, and W. Yu, "Multi-cell MIMO cooperative networks: A new look at interference," *IEEE journal on selected areas in communications*, vol. 28, no. 9, pp. 1380–1408, 2010.
- [8] H. V. Balan, R. Rogalin, A. Michaloliakos, K. Psounis, and G. Caire, "Achieving high data rates in a distributed MIMO system," in *Proceedings of the 18th annual international conference on Mobile computing and networking*, 2012, pp. 41–52.
- [9] S. Gollakota, S. D. Perli, and D. Katabi, "Interference alignment and cancellation," in *Proceedings of the ACM SIGCOMM 2009 conference on Data communication*, 2009, pp. 159–170.
- [10] K. C.-J. Lin, S. Gollakota, and D. Katabi, "Random access heterogeneous MIMO networks," *ACM SIGCOMM Computer Communication Review*, vol. 41, no. 4, pp. 146–157, 2011.
- [11] W. Roh and A. Paulraj, "MIMO channel capacity for the distributed antenna," in *Proceedings IEEE 56th Vehicular Technology Conference*, vol. 2. IEEE, 2002, pp. 706–709.
- [12] W. Choi and J. G. Andrews, "Downlink performance and capacity of distributed antenna systems in a multi-cell environment," *IEEE transactions on wireless communications*, vol. 6, no. 1, pp. 69–73, 2007.
- [13] A. D. Wyner, "Shannon-theoretic approach to a Gaussian cellular multiple-access channel," *IEEE Transactions on Information Theory*, vol. 40, no. 6, pp. 1713–1727, 1994.
- [14] J. Choi, "Compressive random access for MTC in distributed input distributed output systems," in *2017 IEEE 85th Vehicular Technology Conference (VTC Spring)*. IEEE, 2017, pp. 1–5.
- [15] A. Davydov, G. Morozov, I. Bolotin, and A. Papathanassiou, "Evaluation of joint transmission CoMP in C-RAN based LTE-A HetNets with large coordination areas," in *2013 IEEE Globecom Workshops (GC Wkshps)*. IEEE, 2013, pp. 801–806.
- [16] A. Barbieri, P. Gaal, S. Geirhofer, T. Ji, D. Malladi, Y. Wei, and F. Xue, "Coordinated downlink multi-point communications in heterogeneous cellular networks," in *2012 Information Theory and Applications Workshop*. IEEE, 2012, pp. 7–16.
- [17] G. Li, S. Zhang, X. Yang, F. Liao, T.-f. Ngai, S. Zhang, and K. Chen, "Architecture of GPP based, scalable, large-scale C-RAN BBU pool," in *2012 IEEE Globecom Workshops*. IEEE, 2012, pp. 267–272.
- [18] B. Clerckx, G. Kim, J. Choi, and Y.-J. Hong, "Explicit vs. implicit feedback for SU and MU-MIMO," in *2010 IEEE Global Telecommunications Conference GLOBECOM 2010*. IEEE, 2010, pp. 1–5.
- [19] H. Q. Ngo, A. Ashikhmin, H. Yang, E. G. Larsson, and T. L. Marzetta, "Cell-free massive MIMO versus small cells," *IEEE Transactions on Wireless Communications*, vol. 16, no. 3, pp. 1834–1850, 2017.
- [20] Ngo, Ashikhmin, Yang, Larsson, and Marzetta, "Cell-free massive MIMO: Uniformly great service for everyone," pp. 201–205, 2015.
- [21] E. Björnson and L. Sanguinetti, "Scalable cell-free massive MIMO systems," *IEEE Transactions on Communications*, vol. 68, no. 7, pp. 4247–4261, 2020.
- [22] E. Nayebi, A. Ashikhmin, T. L. Marzetta, H. Yang, and B. D. Rao, "Precoding and power optimization in cell-free massive MIMO systems," *IEEE Transactions on Wireless Communications*, vol. 16, no. 7, pp. 4445–4459, 2017.
- [23] H. Q. Ngo, L.-N. Tran, T. Q. Duong, M. Matthaiou, and E. G. Larsson, "On the total energy efficiency of cell-free massive MIMO," *IEEE Transactions on Green Communications and Networking*, vol. 2, no. 1, pp. 25–39, 2017.
- [24] E. Björnson and L. Sanguinetti, "Making cell-free massive MIMO competitive with MMSE processing and centralized implementation," *IEEE Transactions on Wireless Communications*, vol. 19, no. 1, pp. 77–90, 2019.
- [25] G. Interdonato, P. Frenger, and E. G. Larsson, "Scalability aspects of cell-free massive MIMO," in *ICC 2019-2019 IEEE International Conference on Communications (ICC)*. IEEE, 2019, pp. 1–6.
- [26] G. Interdonato, H. Q. Ngo, E. G. Larsson, and P. Frenger, "How much do downlink pilots improve cell-free massive MIMO?" in *2016 IEEE Global Communications Conference (GLOBECOM)*. IEEE, 2016, pp. 1–7.
- [27] Y. Zhang, B. Di, H. Zhang, J. Lin, C. Xu, D. Zhang, Y. Li, and L. Song, "Beyond cell-free MIMO: energy efficient reconfigurable intelligent surface aided cell-free MIMO communications," *IEEE Transactions on Cognitive Communications and Networking*, vol. 7, no. 2, pp. 412–426, 2021.
- [28] E. Björnson and L. Sanguinetti, "A new look at cell-free massive MIMO: Making it practical with dynamic cooperation," in *2019 IEEE 30th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*. IEEE, 2019, pp. 1–6.
- [29] A. Forenza, S. Perlman, F. Saibi, M. Di Dio, R. Van Der Laan, and G. Caire, "Achieving large multiplexing gain in distributed antenna systems via cooperation with pcell technology," in *2015 49th Asilomar Conference on Signals, Systems and Computers*. IEEE, 2015, pp. 286–293.
- [30] J. Qiu, K. Xu, X. Xia, Z. Shen, and W. Xie, "Downlink power optimization for cell-free massive MIMO over spatially correlated Rayleigh fading channels," *IEEE Access*, vol. 8, pp. 56214–56227, 2020.
- [31] J. Qiu, K. Xu, Z. Shen, W. Xie, D. Zhang, and X. Li, "Downlink performance analysis of cell-free massive MIMO over spatially correlated Rayleigh channels," in *2019 IEEE 19th International Conference on Communication Technology (ICCT)*. IEEE, 2019, pp. 122–127.
- [32] G. Interdonato, H. Q. Ngo, E. G. Larsson, and P. Frenger, "On the performance of cell-free massive MIMO with short-term power constraints," in *2016 IEEE 21st International Workshop on Computer Aided Modelling and Design of Communication Links and Networks (CAMAD)*. IEEE, 2016, pp. 225–230.
- [33] J. P. Lemayian and J. M. Hamamreh, "Novel small-scale NOMA communication technique using auxiliary signal superposition," in *2020 International Conference on UK-China Emerging Technologies (UCET)*. IEEE, 2020, pp. 1–4.
- [34] J. Lemayian and J. Hamamreh, "Hybrid MIMO: a new transmission method for simultaneously achieving spatial multiplexing and diversity gains in MIMO systems," *RS Open Journal on Innovative Communication Technologies*, vol. 2, no. 4, 2021.
- [35] J. M. Hamamreh, E. Basar, and H. Arslan, "OFDM-subcarrier index selection for enhancing security and reliability of 5G URLLC services," *IEEE Access*, vol. 5, pp. 25 863–25 875, 2017.
- [36] M. Abewa and J. M. Hamamreh, "Multi-user auxiliary signal superposition transmission (mu-as-st) for secure and low-complexity multiple access communications," *RS Open Journal on Innovative Communication Technologies*, vol. 2, no. 4, 6 2021.
- [37] I. S. Gradshteyn and I. M. Ryzhik, *Table of integrals, series, and products*. Academic press, 2014.
- [38] S. Khalid, S. Iqbal, and J. M. Hamamreh, "A new enhanced coordinated multi-point (CoMP) transmission design for cancelling inter-cell interference for cell-edge users using superimposed supporting signals," *RS Open Journal on Innovative Communication Technologies*, vol. 2, no. 6, p. 12, 2021.
- [39] J. M. Hamamreh, M. Abewa, and J. Lemayian, "New non-orthogonal transmission schemes for achieving highly efficient, reliable, and secure multi-user communications," *RS Open Journal on Innovative Communication Technologies*, vol. 1, no. 2, p. 12, 2020.
- [40] S. Iqbal, S. Khalid, and J. M. Hamamreh, "A new MIMO technique utilizing superimposed auxiliary signals for simultaneously achieving spatial multiplexing and diversity gains in MIMO-Aided communication systems," *RS Open Journal on Innovative Communication Technologies*, vol. 2, no. 4, p. 6, 2021.
- [41] J. M. Hamamreh and J. P. Lemayian, "A novel small-scale nonorthogonal communication technique using auxiliary signal superposition with enhanced security for future wireless networks," 2020.
- [42] S. Iqbal, S. Khalid, and J. M. Hamamreh, "Improving the performance of cell-edge users in 6G and beyond networks by utilizing a novel precoding-based hybrid CoMP transmission design,"
- [43] M. Abewa and J. M. Hamamreh, "D-SEAD: A novel multi-access multi-dimensional transmission technique for doubling the spectral efficiency per area and per device."
- [44] M. Kirik and J. M. Hamamreh, "Interference signal superposition-aided MIMO with antenna number modulation and adaptive antenna selection for achieving perfect secrecy," *RS Open J. Innov. Commun. Technol*, vol. 2, pp. 1–11, 2021.

- [45] S. Iqbal and J. M. Hamamreh, "Improving throughput and reliability performance of future 6G-IoT communication systems using signal superposition-based dual transmission," *RS Open Journal on Innovative Communication Technologies*, vol. 2, no. 5, p. 9, 2021.
- [46] J. P. Lemayian and J. M. Hamamreh, "Physical layer security analysis of hybrid MIMO technology," *RS Open Journal on Innovative Communication Technologies*, vol. 2, no. 5, p. 8, 2021.
- [47] F. Z. Muhammad, H. M. Furqan, and J. M. Hamamreh, "Multi-cell, multi-user, and multi-carrier secure communication using non-orthogonal signals' superposition with dual-transmission for IoT in 6G and beyond," 2021.
- [48] M. F. Zia and J. M. Hamamreh, "An advanced non-orthogonal multiple access security technique for future wireless communication networks," *RS Open J. Innov. Commun. Technol.*, vol. 1, no. 2, 2020.
- [49] J. M. Hamamreh and H. Arslan, "Joint PHY/MAC layer security design using ARQ with MRC and null-space independent PAPR-aware artificial noise in SISO systems," *IEEE Transactions on Wireless Communications*, vol. 17, no. 9, pp. 6190–6204, 2018.
- [50] J. M. Hamamreh, M. Yusuf, T. Baykas, and H. Arslan, "Cross MAC/PHY layer security design using ARQ with MRC and adaptive modulation," in *2016 IEEE Wireless Communications and Networking Conference*. IEEE, 2016, pp. 1–7.
- [51] J. M. Hamamreh, E. Guvenkaya, T. Baykas, and H. Arslan, "A practical physical-layer security method for precoded OSTBC-based systems," in *2016 IEEE Wireless Communications and Networking Conference*. IEEE, 2016, pp. 1–6.
- [52] M. Kirk and J. M. Hamamreh, "A novel interference signal superposition algorithm for providing secrecy to subcarrier number modulation-based orthogonal frequency division multiplexing systems," *Transactions on Emerging Telecommunications Technologies*, p. e4678, 2022.



SADIQ IQBAL in 2016, obtained his electrical engineering B.E. degree from QUEST University Nawabshah, Pakistan. He is presently pursuing the master's (M.Sc.) degree in electrical and computer engineering. He is currently with Antalya Bilim University, Turkey. His research interests include physical layer technology, 5G communication networks, artificial intelligence, IoT, and its applications.



Jihad M. Hamamreh is the Founder and Director of WISLAB (wislabi.com), and A. Professor at the Electrical and Electronics Engineering Department, Antalya Bilim University. He received his Ph.D. degree in telecommunication engineering and cyber systems from Istanbul Medipol University, Turkey, in 2018. Previously, he worked as a Researcher at the Department of Electrical and Computer Engineering at Texas AM University. He is the inventor of more than 20+ patents and an author of more than 85+ peer-reviewed scientific papers along with several

book chapters. His innovative patented works won gold, silver, and bronze medals at numerous international invention contests and fairs. His current research interests include wireless physical and MAC layer security, orthogonal frequency-division multiplexing and multiple-input multiple-output systems, advanced waveform design, multidimensional modulation techniques, and orthogonal/non-orthogonal multiple access schemes for future wireless systems. He is a serial referee for various scientific journals as well as a TPC member for several international conferences. He is an Editor at Researcherstore, RS-OJICT journal, and Frontiers in Communications and Networks. Email: jihad.hamamreh@gmail.com.

IoT and XBee Based Central Car Parking Management System

Seda Ustun Ercan and Mohammed Sufyan Mohammed

Abstract— Parking in congested places frequently leads to major issues like air pollution, traffic jams, etc. This paper presents a prototype for an Internet of Things (IoT) based central car parking management system that attempts to minimize car parking problems. The proposed system is relatively scalable and consumes less power compared to other parking systems. This system uses Arduino microcontroller-based TCRT5000 infrared sensors to detect vehicles in the parking spaces. XBee peripheral modules are located in different parking lots at different locations to transmit the sensed data to XBee central module that is placed close to the parking lots. The central node then updates all data to the Blynk IoT platform via Wi-Fi wireless technology present in the NodeMCU board. The availability of parking lots can be viewed using a smartphone through Blynk application.

Index Terms— Blynk platform, central parking system, Internet of Things, NodeMCU, XBee.

I. INTRODUCTION

PARKING OF vehicles in the parking area is becoming a hard task as the number of vehicles raises while the amount of parking spaces is limited. The outcome would be that the drivers would expend some time looking for parking places, slowing down vehicle flow and increasing traffic congestion. Because parking information is not readily available to drivers seeking for available parking spaces, there is a situation where traffic is backed up in parking lots and hunting for parking places. [1].


Different potential measures have been used to develop car parking management systems such as building new parking facilities which is rather expensive or designing automated parking system where mechanical or hydraulic car lifts regulate vehicles in the parking spaces [2]. An alternative to these very

costly approaches is to augment existing parking facilities with smart systems that can allow more efficient use of parking facilities, as well as allowing the drivers to easily and quickly find available parking space. For example, use green/red light marks to indicate if parking space is currently occupied or free for parking [3]. This solution can be used in dark areas such as parking lots in shopping malls or company buildings, where these marks can be spotted. Such a solution faces some difficulties in exterior areas, like daylight causes the drivers not to recognize these marks light. Therefore, the detecting sensors are commonly connected to display screens or navigation boards, to indicate parking occupancy information [4]. However, this kind of systems are closed and do not communicate with each other, otherwise they can make finding parking spaces much easier. With the growth of Internet of Things (IoT) technology, finding available parking spaces is becoming much easier [5].


There are a lot of cheap and low-energy devices that create IoT networks. These devices have the ability to communicate with each other and provide helpful information on available parking spaces in a citywide manner. With such useful information, the IoT based parking systems could provide various necessary statistics, not only for the drivers and parking managers but also for municipalities or city planners. Several different methods can be used for communication between IoT devices. A wired connection could be an issue in external areas, as it is required to connect all the relevant devices. For this reason, a wireless connectivity makes the solution convenient and more scalable. The most common wireless communication technologies, such as Bluetooth or Wi-Fi, consume a lot of power and they are not appropriate for battery-powered devices. The recently launched zigbee technology employs the IEEE 802.15.4 standard and is a revolutionary, all-encompassing wireless communication technology [6]. When compared to other wireless networks like Wi-Fi and Bluetooth, Zigbee is one of the finest technologies to employ because it has many features like a very low power consumption, less expensive, simplicity to develop and deploy, support of large number of nodes, strong security, and high data dependability additionally, it has self-healing, self-routing, and fault tolerance capabilities [7]. It supports point to point and point to multipoint topology and also supports mesh network topology.

Different smart car parking management systems [8-13] have been suggested which utilize various modules like Arduino ethernet microcontroller or Wi-Fi based boards to transfer the

SEDA USTUN ERCAN, is with Department of Electrical-Electronics Engineering, Ondokuz Mayıs University, Samsun, Turkey, (e-mail: sedau@omu.edu.tr).

 <https://orcid.org/0000-0002-8688-5852>

MOHAMMED SUFYAN MOHAMMED, is with Department of Electrical-Electronics Engineering, Ondokuz Mayıs University, Samsun, Turkey, (e-mail: mohammedsufyan1986@gmail.com).

 <https://orcid.org/0000-0002-6299-9003>

Manuscript received August 11, 2022; accepted December 28, 2022.

DOI: [10.17694/bajece.1160300](https://doi.org/10.17694/bajece.1160300)

sensed data into the cloud. These Arduino or Wi-Fi boards are placed in each parking lot. This certainly increases the cost due to the necessity to provide an internet connection for each Wi-Fi board. In addition to, it raises the power consumed in general. In the proposed parking system, instead of Wi-Fi module, we keep XBee module at each parking lot, and one as a central node that gathers all the sensed data and transfers it into the cloud using Wi-Fi wireless technology. Only the central node needs an internet connection to transmit the information to the cloud. So, the cost and power consumption are reduced. In another project, a real-time system that enables drivers to effectively discover and limit the vacant parking spaces remotely through website and Android App is designed and deployed [14].

The fundamental benefit of mesh networks is that other nodes may continue to communicate without interruption even if one of the links breaks. The hardware that transmits the Zigbee communication protocol is called XBee, and it is made by the digi international firm [15]. As seen in Figure 1, the XBee module. It often serves as both a transmitter and a receiver in wireless communications. XBee modules typically operate at 2.4 GHz with a 250 Kbps speed and a variety of antenna types that allow them to span a range of distances. Although the data rate of XBee module is lower than Wi-Fi, it has a wide application where the transmitted data is less and the power consumption is an issue of concern. The XBee modules contain sleep modes to increase battery life, allowing them to operate over several years on low-cost batteries. Connecting the XBee shield to the Arduino microcontroller is required.



Fig.1. XBee module

In this paper, we designed and implemented a prototype of a central car parking management system based on low power XBee wireless devices and IoT wireless technology. The proposed parking system assists drivers to find vacant parking spaces with the reservation facility. Unlike other studies, for the first time, the proposed parking system is usable in both external as well as internal parking facilities and thus the that system could solve all the previously mentioned problems to some extent.

II. SYSTEM OVERVIEW

Several requirements have been identified for the proposed parking system such as : (1) Parking lots occupancy detection – the proposed system must be able to detect the occupancy of individual parking lots. This requirement was achieved by choosing an appropriate set of sensors, specifically infrared sensors. (2) Visualization of available parking lots – the

proposed system should notify the driver of the state of parking lots in a visible manner. This requirement was fulfilled by the development of a graphical user interface (GUI) on Blynk application (app), which presents information related to the parking lots for the drivers. (3) Wireless communication, (4) Low power consumption – to provide scalability and minimize cabling challenges and infrastructure costs. The third requirement was achieved by using Zigbee low power as a wireless communication technology that connects all system nodes. For the fourth requirement, it is largely related to the third requirement, where the low power zigbee technology was used and the selected devices are low power, simple and could run on batteries for a sufficient period of time. (5) Readily available software and components – the proposed parking system must be based on free or open-source software without license fees and well-known or widely used devices. The proposed system devices or components meet the fifth requirement as they have been selected, for which plenty of extensions, user guides, software and libraries are freely available. An overview of the proposed parking system architecture is illustrated in Figure 2. There are two kinds of nodes in the proposed parking system: Peripheral nodes and Central node.

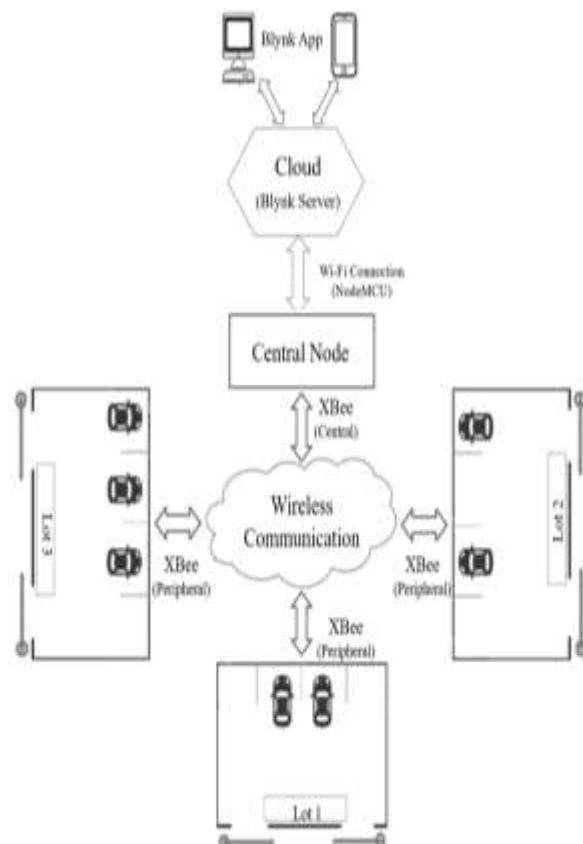


Fig.2. The proposed parking system architecture

Peripheral nodes are located at each parking lot to detect the presence of the vehicles in that area and to report the status to the central node constantly. At each peripheral node, there are Arduino Mega microcontroller, TCRT5000 infrared (IR) sensors, servo motors, liquid crystal display (LCD), matrix keypad, barrier gates and XBee module. Arduino [16] is a common microcontroller, which has many versions. It is

effective for gathering data from different sensors and has plenty of extension shields that can connect it to other devices.

For the proposed parking system, Arduino Mega 2560 development was used, which is cheap, easy to use, and has large storage capacity and low power consumption. It can be easily extended using various shields, such as XBee shield that was used in the proposed parking system. The Arduino microcontroller was programmed using the free and open-source Arduino software IDE, which makes it simple to create code and upload it to the board. The total number of parking spaces occupied in the parking lot was shown on an LCD. Check-in and check-out were controlled by barrier gates, which were driven by servo motors.. Matrix keypad to key in the passwords was used. IR sensors were used to detect the presence of the vehicles at the entrance and exit of the parking lot, and also were used to detect the state of each space within the parking lot. The IR sensors detect if the parking lot is available or not and are directly connected to the corresponding peripheral node, represented by the Arduino Mega microcontroller. A single peripheral node can handle multiple parking spaces at the same time. The IR sensors measure the distance of the existing vehicle. If the measured distance is less than the specified threshold value, then the parking space is identified as occupied. Flags were set for each parking space (flag=1 if occupied, flag=0 if available). To distinguish parking spaces information in each parking lot, the peripheral node sends the flag values along with the name of the parking lot (lot, lot2, etc) to the central node which transfers the sensed data to the cloud server. The central node acts as a central controller that records peripheral nodes data and controls the overall system. This node requires higher computing energy as well as some memory space. For these reasons, the Arduino Mega microcontroller was used as a central node as well.

XBee modules were used to provide communication and information exchange between the central node and peripheral nodes. XBee at the central side receives the data from multiple XBee modules at the peripheral side simultaneously. The peripheral XBee is installed in different parking lots to transfer the sensed information to the central XBee without the need to use internet or Wi-Fi connection at every parking lot. In other words, Wi-Fi protocol is used only at the central node to transfer the received data to the cloud. XBee modules could communicate with each other wirelessly and it could exchange the information over a wide distance up to 2 km range in this case. The central XBee should be at that range to receive the information from multiple peripherals XBee. For illustration, if there are three parking lots, instead of three Wi-Fi modules, we keep three XBee peripherals and one XBee central. Then we specify a point which is roughly within the specified range from the three parking lots and we install the central XBee at that point. XCTU software was used for the configuring of XBee modules. The central XBee feeds the gathered data to the central node which transfers it to the back endless cloud based Blynk server using Wi-Fi connection through the internet. The cloud server acts as a database to store all the necessary records related to the parking lots and the drivers that have access to the parking system. NodeMCU module was used in the proposed parking system to connect the central node with Blynk server. NodeMCU is an open source, simple, programmable and Wi-Fi

enabled module designed for IoT purposes. It includes firmware that runs on the ESP8266 Wi-Fi system on chip. The ESP8266 is a Wi-Fi chip with microcontroller capability and full TCP/IP protocol stack that allows any microcontroller to access the Wi-Fi network. There are many kinds of ESP8266 modules, but for this proposed parking system the ESP12-E NodeMCU module was used and was programmed through Arduino IDE software as well [17]. The implementation of the proposed central and peripheral nodes on a breadboard is illustrated in Figure 3 and Figure 4.



Fig. 3. Proposed central node hardware



Fig. 4. Proposed peripheral node hardware

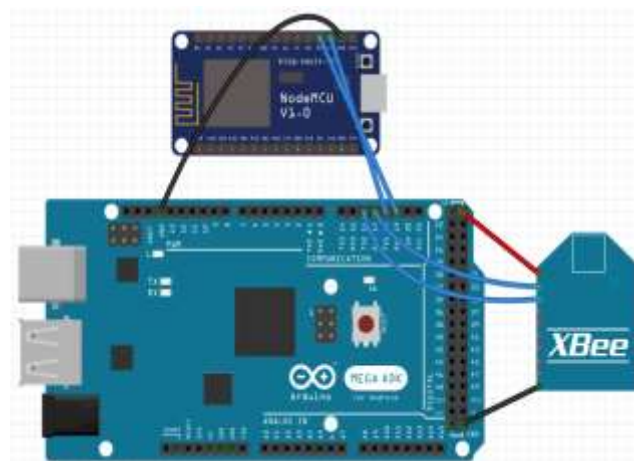


Fig. 5. The proposed central node diagram on fritzing

The interconnection between different electronic components of the proposed central and peripheral nodes are illustrated in Figure 5 and Figure 6. It was designed using fritzing environment. It is worth noting also that the system nodes are powered from separate power sources.

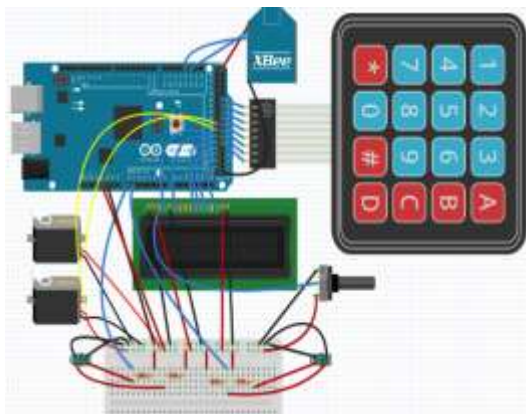


Fig. 6. The proposed peripheral node diagram on fritzing

III. BLYNK INTERNET OF THINGS PLATFORM

IoT technology refers to anything physical that is connected to the internet or exchanging data or information between physical devices and the internet. Blynk is a platform was designed for IoT purposes. It is compatible with both iOS and Android apps and can interact with various microcontrollers such as Arduino, Raspberry Pi, NodeMCU etc. It is able to remotely control devices over the internet and store, display and visualize data from different sensors for any desirable system. There are three main components in the Blynk IoT platform: (1) Blynk Application – It allows creation of graphical user interfaces for any project using different widgets that are provided. (2) Blynk Server – It is an open-source cloud based server, could easily handle thousands of devices and is responsible for all the communications between the smartphone (Blynk app) and the selected hardware. (3) Blynk Libraries – It enables communication, for all the popular hardware platforms, with the server and processes all the incoming and outgoing commands [18].

For the proposed parking system in this paper, a simple GUI was designed and developed using Blynk app, which displays parking lots names along with the total number of available parking spaces in each parking lot. Step-by-step procedures in designing the GUI Blynk app are illustrated in Figure 7. Firstly, we need to download and install Blynk app for android at Google plays store or iOS at app store. For the proposed parking system, two pieces of Value Display tool from Widget Box were used, each one to display the number of vacant parking spaces within each parking lots, as shown in Figure 7 (a). The Value Display tool displays incoming data from sensors or Virtual Pins. By clicking on Value Display tool for a one time, we can prepare it, as shown in Figure 7 (b). One of the Value display tools was labeled as Parking Lot A while the other one was labeled as Parking Lot B, as shown in Figure 7 (c). It can be observed that the two value display tools are related to the same SOURCE which is the Central Car Parking. Then from Input/PIN/Select pin we need to choose Virtual and from PIN we need to choose a virtual port number to provide the

communication with the selected device. Virtual pins are a way to exchange any value or data between Blynk app and the devices. It could sent the values as strings and there are no practical limits on it. The process of sending any value or data from the selected device to Blynk app is done through a function called Blynk. VirtualWrite(Virtual Pin Number, Value) function, while any value or data from Blynk app can be sent to the selected device using Blynk_Write (Virtual Pin virtual pin's number as a parameter in the functions. Also, param.asInt () must be called inside the function, which returns an integer (It could be boolean or byte as well) value, that must be processed to execute other sentences in the main code. In the proposed parking system, V7 virtual port was used for parking lot A and V8 was used for parking lot B, as illustrated in Figure 7 (d) and Figure 7 (e) respectively. So, when XBee central receive data from XBee peripheral, it feeds it to the Arduino Mega at the central node for processing. If the received data is related to the parking lot A, the central node transfers the processed data to the Blynk app through virtual port V7, and if the data is related to the parking lot B, the central node sends it through virtual port V8 to the Blynk app. The designed GUI on Blynk app dashboard with the Value Display tools is shown in Figure 7 (f). And with this the Value Display tools settings have been ended.

The designed GUI on Blynk app also includes parking space reservation option by clicking the reservation button in the designed GUI associated with the desired parking lot. For reservation process purpose the Styled button from Widget Box was used. The Styled button operates in switch or push modes. It allows sending any number value on button press and button release events. By default, the Styled button uses 1/0 (HIGH/LOW) values. The button sends 1 (HIGH) on press and sends 0 (LOW) on release. A title for ON/OFF LABEL was given and V9 virtual pin was used for the Styled button, as shown in Figure 7 (g). It can be seen that the Central Car Parking is the TARGET. After the reservation request is created, the boolean data (low or high) travels to the central node through Blynk server and NodeMCU module. The Arduino microcontroller then processes the received data and begins start to check whether there is any vacancy in the desired parking lot. In the case of the desired parking lot is not available means the central node will send a notification or a statement to the Blynk app dashboard telling that there is no available parking space. In the case of the desired parking lot is available means the central node will send a reservation acknowledgment message along with the entry password to the respective parking lot using XBee modules. The total number of occupied parking spaces at the desired parking lot will increase by one.

At the same time, the central node will update the desired parking lot information and the entry password to the Blynk app. A timer will be started automatically for that reservation. Before the timer expires the driver has to arrive at the desired parking lot and enter the entry password, otherwise the reservation process will get expire and in this case, the information of that desired parking lot will be updated and the updated information will be sent to the central node using XBee modules as well. The central node will be updated the information of the desired parking lot to the Blynk app again. If the driver arrives at the specified time means, the driver has to

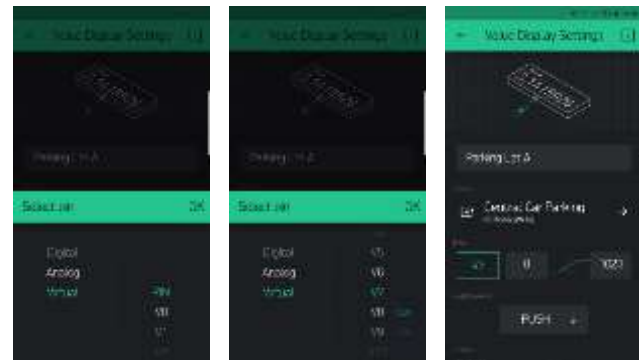
enter the corrected password in order to access the desired parking lot. If the entered password is correct, the barrier gate will get open and allow the driver to enter parking lot area. We have used the labeled Value tool from Widget Box for displaying the entry password and reservation statement. The Labeled Value tool displays incoming data from Virtual Pins or different sensors. It is a better version of the Value Display tool because it contains a formatting string. It was named as “Reservation” and V10 was used as a virtual pin, as shown in Figure 7 (h). It can be seen that the Central Car Parking is the SOURCE, and the virtual pins V7, V8 and V9 are already in use. The designed GUI on Blynk app dashboard with the tools used is shown in Figure 7 (i) where it can be seen all the tools with their names and their virtual ports.

There are a few changes that must be made to the main code or program before uploading. At first, to interface with Arduino software, Blynk library needs to be installed at the library folder of Arduino software IDE. Also, we need to use or to pre-install some other libraries like `#include<ESP8266 Wi-Fi.h>` library, `#include <SPI.h>` library, `#include <SimpleTimer.h>` etc, to programing using Arduino IDE. Node MCU 1.0 (ESP 12-E Module) module has to be selected as the board type. The COM port has to be chosen after connecting the selected device to the computer. The baud rate also must be set up to monitor the output of different devices in the serial monitor of Arduino IDE.

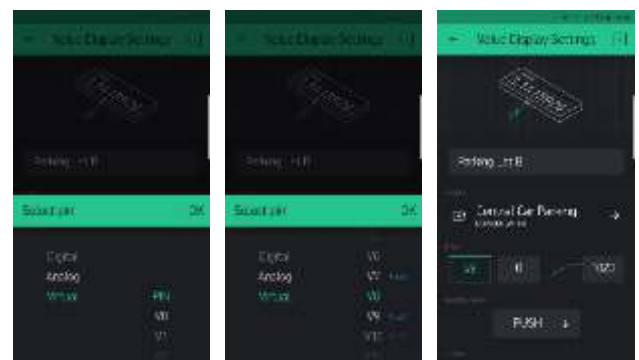
The Wi-Fi SSID and the Wi-Fi password should be changed by the actual Wi-Fi SSID and Wi-Fi password. The Auth Token also needs to be modified in the source code as it is given in the Blynk account. After checking all the connections, adding all the required library files to the Arduino IDE, setting the right ports of the computer and after completing all the steps, the main code has to be compiled and then uploaded to the selected module. After compiling and uploading the Arduino code, the peripheral node starts sending the data from IR sensors to the central node using XBee modules. The central node transfers the data to the Blynk cloud server and Blynk app through Wi-Fi connection. Some of the results of the designed parking system obtained on Blynk app are shown in Figure 8 (a) (b) (c). As an illustration, fig 8 (a) indicates that there are three parking spaces available in the parking lot B while there are only one parking space is available in the parking lot A.



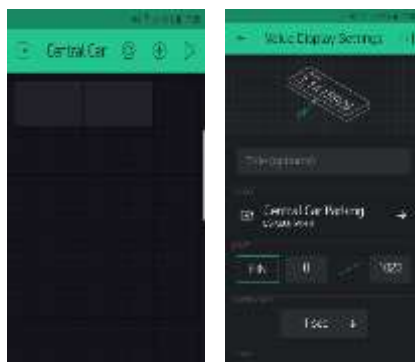
(c)



(d)

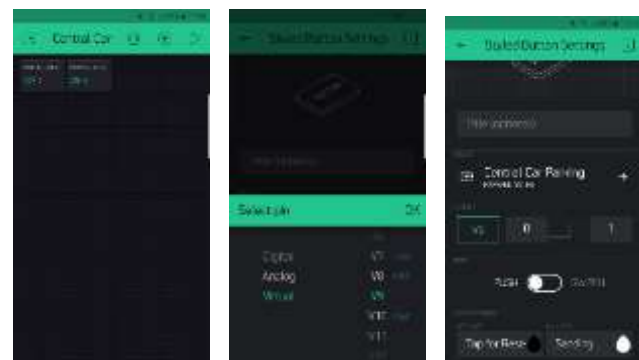


(e)



(a)

(b)



(f)

(g)

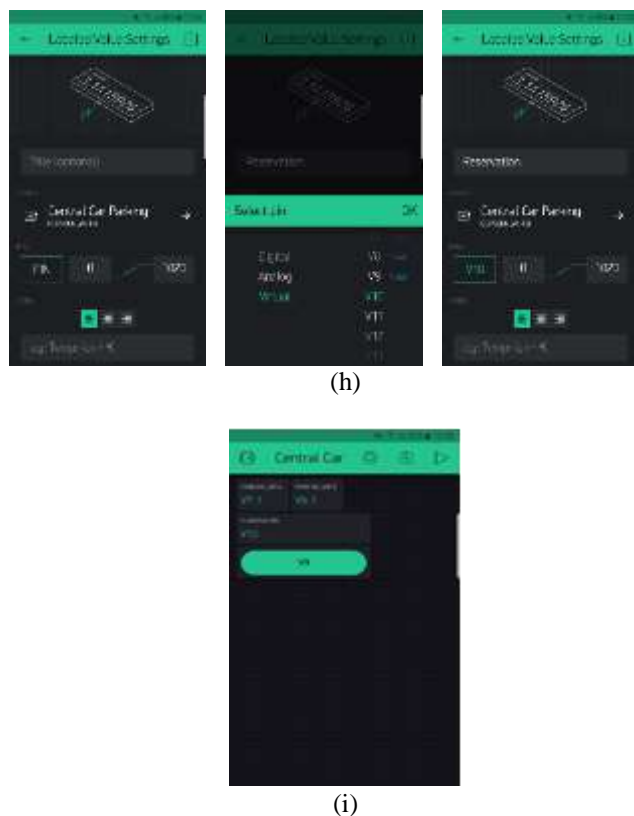


Fig. 7. Illustrates the design stages of the GUI on Blynk app used in this paper

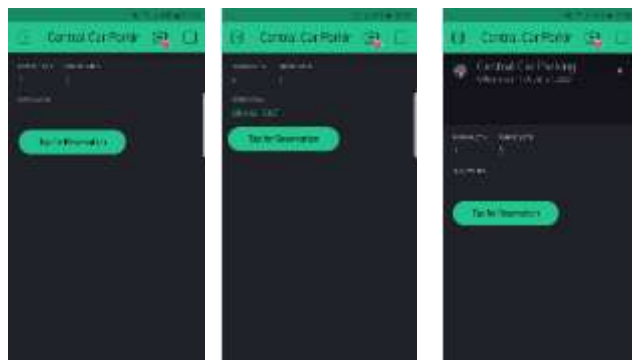


Fig 8. Blynk app showing availability status of parking lots

IV. CONCLUSION

The main purpose of this study is to suggest solutions to the problem of looking for available parking spaces in urban areas, particularly by considering issues such as communication approaches, power consumption and financial cost. More specifically, a prototype of a central car parking management system was designed and implemented based on IoT technology. The proposed parking system is capable of detecting available parking lots for the drivers with high accuracy. It can be applied for the street-based parking lots, as well as interior and exterior parking lots. Moreover, the proposed parking system can be combined with the present IoT based parking systems to provide information about managing parking lots in citywide.

One of the main contributions of the suggested parking system is that it is based on widely available devices utilizing well-known software and components that do not charge license fees. The selected devices make the proposed system scalable and easily deployable as each component can be replaced easily. In the tested prototype of the proposed parking system, Arduino Mega microcontroller was used which appeared to be cost effective and compatible with XBee modules with easy installation or programming and maintenance.

The prototype of the proposed parking system was implemented and tested for various scenarios and accurate results were obtained. The results of the intensive experiments proved that the proposed system successfully detects the presence of the vehicle at each parking lot, and updates the sensed data to the central node and Blynk app effectively. Based on that we conclude that the successful implementation of a central car parking management system can significantly reduce the problem of fuel, time and effort wastage, and pollution caused by blind searching and traffic congestion to a considerable amount.

However, the designed GUI on Blynk app needs to be improved as the prototypes were designed and developed for testing purposes only, and it is not suitable for commercial use. Other IoT based parking systems provide the ability to reserve parking spaces extensively and to pay the fee using a smartphone. In further work, we would develop an application for Android and iOS and test its usability in the real-time environment. Also, the integration of the navigation function into the proposed system is a possible extension in the future work.

REFERENCES

- [1] L. Anjari and A. H. S. Budi, "The development of smart parking system based on NodeMCU 1.0 using the internet of things," IOP Conference Series: Materials Science and Engineering, International Symposium on Materials and Electrical Engineering (ISMEE), Bandung, Indonesia, 16 November 2017.
- [2] T. Bhanusri and K. P. Rao, "Advanced Car Parking System with GSM Supported Slot Messenger," *Journal of Electronics and Communication Engineering (IOSR-JECE)*, vol. 10, issue 1, pp. 14-18, February 2015.
- [3] A. Kianpisheh, N. Mustaffa, P. Limtrairut, and P. Keikhosrokiani, "Smart Parking System (SPS) Architecture Using Ultrasonic Detector," *International Journal of Software Engineering and Its Applications*, vol. 6, no. 3, pp. 55-58, July 2012.
- [4] P. B. Natarajan and S. K. Ghosh, "Design and Implementation of Smart Car Parking System Using Lab VIEW," *International Journal of Pure and Applied Mathematics*, vol. 120, no. 6, October 2018, pp. 329-338.
- [5] A. Khanna and R. Anand "IoT based smart parking system," IEEE 2016 International Conference on Internet of Things and Applications (IOTA), Pune, India, 2016, pp. 266-270.
- [6] H. C. Yee and Y. Rahayu, "Monitoring parking space availability via ZigBee technology," *International Journal of Future Computer and Communication*, vol. 3, no. 6, pp. 377-380, December 2014.
- [7] A. Asaduzzaman K. K. Chidella and M. F. Mridha, "A time and energy efficient parking system using Zigbee communication protocol," IEEE Southeast Con 2015, Fort Lauderdale, FL, USA, 2015, pp. 1-5.
- [8] P. Sadhukhan, "An IoT-based E-parking system for smart cities," IEEE 2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI), Udupi, India, 2017, pp. 1062-1066.
- [9] A. Somani S. Periwal, K. Patel and P. Gaikwad, "Cross Platform Smart Reservation Based Parking System," IEEE 2018 International Conference on Smart City and Emerging Technology (ICSCET), Mumbai, 2018, pp. 1-5.

- [10] B. M. Mahendra, S. Sonoli, N. Bhat, Raju and T. Raghu, "IoT based sensor enabled smart car parking for advanced driver assistance system," 2017 2nd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT), Bangalore, 2017, pp. 2188-2193.
- [11] A. Gupta, P. Rastogi and S. Jain, "Smart Parking System using Cloud based Computation and Raspberry Pi," 2018 2nd International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) India, 2018, pp. 94-99.
- [12] S. Mendiratta, D. Dey and D. Rani Sona, "Automatic car parking system with visual indicator along with IoT," 2017 International conference on Microelectronic Devices, Circuits and Systems (ICMDCS), Vellore, 2017, pp. 1-3.
- [13] G. Aravindh and M. A. Kumar, "An Efficient Car Parking Management System using raspberry-pi," 2020 3rd International Conference on Intelligent Sustainable Systems (ICISS), 2020, pp. 154-158, doi: 10.1109/ICISS49785.2020.9316049.
- [14] A. Anand, A. Kumar, A. N. M. Rao, A. Ankesh and A. Raj, "Smart Parking System (S-Park) – A Novel Application to Provide Real-Time Parking Solution," 2020 *Third International Conference on Multimedia Processing, Communication & Information Technology (MPCIT)*, 2020, pp. 93-96, doi: 10.1109/MPCIT51588.2020.9350429.
- [15] <https://www.digi.com/products/embedded-systems/digi-xbee/digi-xbee-tools/xctu>
- [16] <https://docs.arduino.cc/static/de1277b5f061164f715cb610870edb1a/A000067-datasheet.pdf>
- [17] https://www.nodemcu.com/index_en.html#fr_54747661d775ef1a3600009e
- [18] https://blynk.cloud/dashboard/register?_ga=2.141173108.732515720.1660131406-1285698961.1660131406

BIOGRAPHIES



SEDA USTUN ERCAN was born in Turkey. She received her B.S, M.Sc. and Ph.D. degrees in electrical and electronic engineering from Ondokuz Mayıs University (OMU), Samsun, Turkey, in 2005, 2009 and 2016, respectively. She is currently an Assistant Professor with the Department of Electrical-Electronics

Engineering in Samsun, Turkey. Her research interests include communication and digital signal processing.



MOHAMMED SUFYAN MOHAMMED was born in Iraq, in 1986. He received B.S and M.Sc. degrees in computer engineering from Mosul University, Iraq in 2010 and electrical and electronic engineering from Ondokuz Mayıs University (OMU), Samsun, Turkey, in 2020 respectively.

LSTM-GRU Based Deep Learning Model with Word2Vec for Transcription Factors in Primates

Ali Burak Oncul


Abstract— The study of the structures of proteins and the relationships of amino acids remains a challenging problem in biology. Although some bioinformatics-based studies provide partial solutions, some major problems remain. At the beginning of these problems are the logic of the sequence of amino acids and the diversity of proteins. Although these variations are biologically detectable, these experiments are costly and time-consuming. Considering that there are many unclassified sequences in the world, it is inevitable that a faster solution must be found. For this reason, we propose a deep learning model to classify transcription factor proteins of primates. Our model has a hybrid structure that uses Recurrent Neural Network (RNN) based Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) networks with Word2Vec preprocessing step. Our model has 97.96% test accuracy, 97.55% precision, 95.26% recall, 96.22% f1-score. Our model was also tested with 5-fold cross-validation and reached 97.42% accuracy result. In the prepared model, LSTM was used in layers with fewer units, and GRU was used in layers with more units, and it was aimed to make the model a model that can be trained and run as quickly as possible. With the added dropout layers, the overfitting problem of the model is prevented.

Index Terms—Protein classification, Hybrid deep learning, Word2Vec, LSTM, GRU

I. INTRODUCTION

THE HEREDITARY material of living things is Deoxyribonucleic Acid (DNA). DNA sequences consist of adenine (A), guanine (G), cytosine (C), and thymine (T) nucleotides and, with their various sequences, form the structures necessary for the creation and survival of living things with all their units [1], [2]. One of these structures is amino acids. There are over 300 amino acids in nature. 20 of these amino acids are mostly found in mammals and plants. These amino acids produced by DNA are arranged in different ways to form proteins. Proteins are involved in almost all vital tasks in living things, such as growth, measures under stress, and communication of cells. [3]. Many amino acid sequences have been discovered in light of scientific studies and advances in science. The discovered sequences belong to many different realms, species, and families.

ALİ BURAK ÖNCÜL, is with Department of Computer Engineering Department of Kastamonu University, Kastamonu, Turkey, (e-mail: boncul@kastamonu.edu.tr).

 <https://orcid.org/0000-0001-9612-1787>

Manuscript received October 18, 2022; accepted Nov 13, 2022.
DOI: [10.17694/bajece.1191009](https://doi.org/10.17694/bajece.1191009)

Each unit that is different has many different structural features. These different features determine the types and groups of these amino acid sequences, namely proteins.



Fig.1. Example of TF protein (bHLH) [4].

Transcription is the first step of gene expression from DNA to RNA transcript production in a gene. This process, together with the continuation steps, carries out the production of the protein. Transcription factors (TFs) are genomic regions that can bind to certain DNA sequences and fragments, directing gene expression in different ways in cells and, thus, organisms. TFs are proteins that manage the transfer of genetic information from DNA to mRNA. Each TF has a structure specific to the sequences [5]–[7]. Animal TF proteins are proteins that have essential roles in regulating many vital functions of cells, such as their development, communication, response under stress, and cell cycle. These proteins, like other proteins, have been identified and sequenced by biological experiments, and new sequences are being discovered every day [8]. Figure 1 shows a TF protein from basic helix-loop-helix TF family.

Conventional biological research methods can classify these many different proteins. However, these classification processes consist of both costly and relatively long-term experiments. In addition, the possibility of error by the experimenter is among the handicaps of these traditional methods [9]. This situation has pushed researchers to search for new analysis methods. In this field, especially in studies conducted jointly with computer scientists, these data were firstly the subject of statistical models [10]. Examples of these models are the studies prepared with the Hidden Markov Model (HMM) [11], which is widely used [12], and the Basic Local Alignment Search Tool (BLAST) [13] study. Acting on the principles of statistical science, these studies act and make predictions according to the probabilities of each amino acid in the positions in the protein sequence. In this way, the classes and structures of the arrays are determined. However, the models developed by these studies work better with annotated, i.e., additional labeled data such as sequence or the primat

name. For this reason, the success rate will be relatively low in plain sequences [14], [15].

There have also been studies in which artificial intelligence applications such as machine learning and deep learning, which are used in different fields and studies, are used in the classification of proteins in the field of bioinformatics. Examples of machine learning algorithms used for this purpose are Naive Bayes (NB) Classifier, Gradient Descent Algorithm (GD), K-Nearest Neighbor (KNN) algorithm, and Support Vector Machine (SVM). Recursive neural networks (RNN) and convolutional neural networks (CNN) based deep learning applications, which have been brought to the literature with the developments in computer science and the latest innovations in artificial intelligence, are also used for different kinds of classification problems in this field [16]. In addition to these studies, studies have been carried out on similar data with various deep learning models, together with the developments in the field [17].

II. RELATED WORKS

Looking at the literature, different bioinformatic studies have been developed in addition to biological studies to analyze various protein sequences. These studies draw their strength from the intersection of biological studies and computer science. Although these may be statistically based, in recent studies, the emphasis has been on artificial neural networks, machine learning, and especially deep learning studies [18]. These studies have been included in the literature with different data sets and models. In a study dealing with gene ontology and protein function analysis, PSI-BLAST and position-specific scoring matrix (PSSM) were used, along with the alignment results of PSI-BLAST and position information of PSSM, protein functions, and annotations were determined in new genomes [19]. A KNN-based OET-KNN model was developed in a study that controls protein sequences to become enzymes and determines their sub-functional classes. This model, with its variations, has achieved success on a scale of 86%-98% [20]. In another study, which detects the enzymatic functions of proteins, the ECPred model was developed, and a combination of Pepstats-SVM, SPMAP, and BLAST-KNN models contributed to the literature [21]. The SVM structure was used

in yet another gene ontology-based model, and a new approach was brought to the field [22]. These studies are mid-term studies, and since some of them require some additional experimentally acquired information besides lean protein sequences, current models that only work with protein sequences and do not require any additional information were needed [15].

The model working with Keras embeddings, ProtVec vector representation, and SVM using the Swiss-Prot dataset from studies conducted with arrays only has a success rate of 93% [23]. Again with the same dataset, Keras embeddings achieved success on a scale of 81.2% to 91.24% with a vector representation tool like ProtVec, model variations prepared with LSTM and CNN [16]. Another study used in enzyme research, using a pre-trained model with the Swiss-Prot dataset, achieved 97% success with the support of CNN and LSTM networks [15]. In this study, the sequences were divided into 3-character parts and processed just like the sentences and words of a text [23]. In another study of transport proteins, an accuracy rate of 85.8% was achieved with PSSMs, CNN, and GRU models [24]. In a study prepared with several different artificial neural network models such as KNN, Naive Bayes, and SVM, classification was made on the mouse protein dataset with Down syndrome [25]. In the study on urease activity in full-fat soybean production, a model was prepared with CNN-LSTM networks, and 96.57% and 90.29% successes were obtained according to the number of classes [26]. In a study prepared for protein homology detection, preprocessing with single-hot coding and model preparation with bidirectional LSTM were performed, and the model achieved 97% success [27]. In another study, the Pfam seed dataset was used, and code dictionary and LSTM were used. In accordance with the nature of the code dictionary structure, a number was assigned for each amino acid in the sequences, and analyzes were made on the digitized sequences [27], [28]. The 3-layer LSTM model used in a transfer learning-based study has brought 85% success [29]. In another protein classifier, the power of residual blocks of the CNN-based ResNet architecture was utilized, and a success rate of 93.7% was achieved [30]. The studies given above have demonstrated the importance of artificial intelligence in studies such as the classification of proteins, detection of binding sites, and detection of proteins from genes [31]–[33].

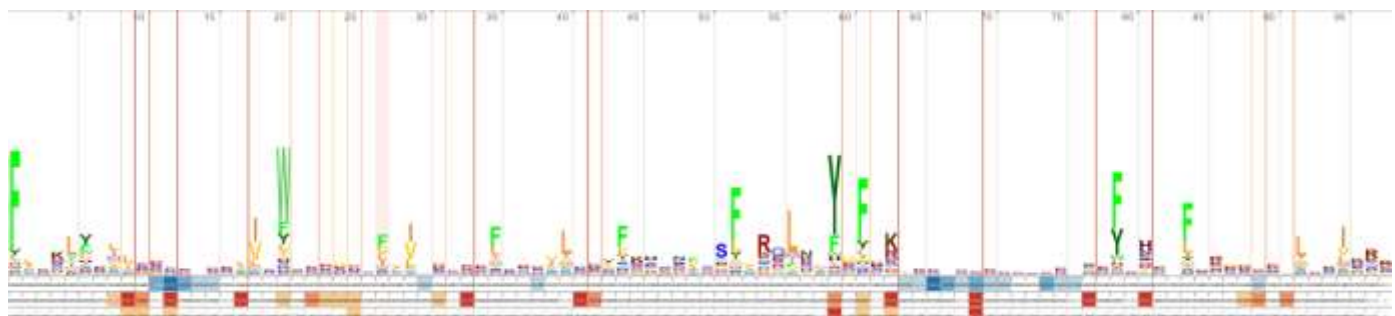


Fig. 2. HSF TF protein family motif structure [34].

III. MATERIALS AND METHODS

Transcription factor proteins in primates, including humans and various monkey species, were obtained from fragmented and

scattered data from the Animal Transcription Factor Database (AnimalTFDB), open to researchers. AnimalTFDB, with its current version, is a database containing TFs and their

cofactors, various annotations and binding domains, and subsections of these TFs in 183 animal genomes. Various analysis and alignment tools are also included in AnimalTFDB. It also includes additional information such as gene expression, post-translational modification and mutation information, and paralog and ortholog information [35].

A. Representation of Protein Sequences

Various proteins are formed by the sequential arrangement of amino acids according to the structural information [36]. This difference in structural information creates the unique structures of different proteins and, thus, different 3D structures [3]. Although each protein formed is different, they form different protein families together with the proteins with which they are structurally common. This structural commonality is called motifs in sequences. Sequences with identical motifs are proteins of the same family [37]. Figure 2 shows an example of a motif belonging to the Heat Shock Factor (HSF) TF protein family.

The letters representing amino acids in the motif in Figure 2 represent the probability of having the amino acid at the letter's position—the larger the representation, the more likely that amino acids will be found at the location. In addition to this information, when the motif is examined, it is seen that there may be many different amino acids in each position of the motif. For this reason, it is almost impossible to know which family each protein belongs to by examining the sequence character representations by human manually. For this reason, biological experiments or various artificial intelligence applications are some of the solutions. Considering that biological experiments are also a long time and costly, artificial intelligence applications stand out as one of the most effective solutions. Figure 3 shows a raw sample sequence of Homo sapiens, downloaded from AnimalTFDB [35].

```
>ENSP00000306549.3 ENSG00000169953 HSFY2 HSF
MAHVSSSETQDVSPKDELTAESTRSPICEHTFPGDSLRSMIEEHAFQVLSQGSLLSESP
SYTVCVSEPKDDDFLSLNFPRKLWKIVESDQFKSISWDENGTCIVINEELFKKEILETK
APYRIFQTDALIKSFVRQLNLYGFSKIQNQFQSAFLATFLSEKESSVLSKLFYNNPNF
KRGYPQLLVVYKRRIGVKNASPISTLFNEDFNKKHFRAGANMENHNSALAAEASEESLFS
ASKNLNMLPTRESSVRQIIANSSVPIRSGFPSPSTSVGPSEQIATDQHAILNLQTTIH
MHSHSTYMQARGHIVNFITTTTSQYHIIISPLQNGYFGLTVEPSAVPTRYPLVSVNEAPYR
NMLPAGNPWLQMPTIADRSAAPHSRRLALQPSPLDKYHPNYN
```

Fig. 3. A raw sample sequence of Homo sapiens from AnimalTFDB [35].

In the representation in Figure 2, the character ">" denotes the beginning of each new sequence. The first line contains the sequence Protein ID, Ensembl ID, gene symbol, and protein family information, respectively. From the second line to the next ">" character, there is the protein sequence. These files for each species are stored separately in AnimalTFDB in FASTA format.

These files in FASTA format, downloaded separately for all primates from AnimalTFDB, are created with the prepared Python scripts and replication operations, with a single line containing only protein families and sequences in a single line, with each protein family - protein sequence pair represented on a single line, containing all primates. combined in a tab-delimited text (.txt) file. As a result, a data set consisting of 72 different TF protein families belonging to 24 primate species

and containing 36242 sequences was created. Figure 4 shows a sample part of this dataset.

```
family sequence
Homoebob MDTSRPGAFVLSSAPLAALHMAENKTSLFYALGQAPGAPALGDLGALPLGTGPHGISDILGRPYTA
AGDGLLGLPRLNGLASSAGVYFQPAAYARDYPLAELPQPPPIFWQYVQSGAPWRPRLACAPAGDVLKDKKKHSRPT
FSGQZIFALEKTEQTKYLASPERARLAYSLGHTESQVKNYNNRRTKMKRHAENASAKKKQDSAEKLVGGSDAEDDD
EYNPLDPSNDDKTSRLKKHKPSNLALYSPCGGAGDAL
ZF-C2H2 MEIKPLPARGLTQGGGGSLFAGGGRVHRSPDPFAGQVPTRRLLLRGPDGGGPRRREAPTAAGRGPSSL
APRPDQSGGGDDFFLVLLDPVGGDVETAGSQAAGPVLREEAAGPQLQGGSSANPAGRPALGPRCLSAVSTPAPISAPF
PAAAFAGTVTITHNDLLRFENGVLTLATPPHAMDPGAAPAGQDCLVAPQAGFPQAAGSGDCELPDOLLAEAEAPFA
PAPEEAEQPAALQPRGRLGSAPOVVLVLCPEAGCQQTFAKKHKLKVHLLTHSSGGQRPFKCPLGGCMTFTTSYKLKRH
LQSHDKLRPFQCPAEGCGKSFITVYNLKAHMKQHEQENSFKCEVCESFPTQAKLSAHQRSHFPERPYQCAFSGCKKTFIT
VSALFSHNRAHFREQLFSCSFPGCSKQYDKACRLXHLRSHTEGRRPFLCOFGGQWNTSMGKLLRHKKHDDRRRTCPV
EGCGKSFTRAEHLKGRHSITHLGTPFVCPVEGCCARFSARSSLYHSHKHLQDVDTKSRCPITCNKILFTSKHSMKTHMAK
RHKVGQDLAQLEAAANSLTPSSELTSGGNDLSEAEIVSLFSDVPDGTSAVLDOTALVNSGILTIDVASVSTLAGNLANN
MNSVGGVDPPLMATSOPPSLDLSLFFGTATGFGQSLDMDEVSVTVGPGLSGLAVKNSSPEPQALTPSSKLTVDIT
DALTPSSTLCENSSELLTPTKAEWVHPDSDFGREGETQGFNPAAGNHSQKETDLVTVTGSCSFLY
```

Fig. 4. A sample part from the prepared data set.

B. Data Preprocessing Techniques

The structures in the form of "MDTSRPGAFVLSSAPLA...", which is the representation of proteins in letters in character representation, is a meaningful structure for researchers in biology and bioinformatics. However, computers can only evaluate these sequences as words or sentences, and it is not possible for them to derive a direct meaning. Just as preprocessing is done on images, such as reducing size, changing color values, and extracting images locally, various operations, such as trimming sentences and texts, clipping suffixes, and adjusting sentences to equal length, are performed in natural language processing studies. Classification of protein sequences, which is a sub-branch of natural language processing studies, also requires a similar set of preprocessing steps [37].

In this study, a series of preprocessing steps, such as adjusting the protein sequences to a fixed length, shortening the long sequences, extending the short sequences by filling from the end, and preparing the embeddings from the sequences, were applied to the prepared data set.

Protein sequences differ from the data used in other natural language processing studies. Normally, sentences in texts consist of words, while protein sequences, each of which can be evaluated as a sentence, are in one piece. For this reason, previous studies have been done in the literature on the representation of each amino acid of the sequence with a number [28], [38]. However, suppose protein sequences are preprocessed at the amino acid level. In that case, achieving the highest desired success will be difficult since the relationships between these amino acids, that is, the relationship of each amino acid with the previous and next amino acid, cannot be revealed.

The probability that amino acids or amino acid groups come before or after each other, that is, their proximity to their location, is essential for capturing motifs and successful protein classification, just as knowing the similarities of words in natural language processing studies [39]. For this reason, it is necessary to know the affinity between these words (k-mers) by separating the protein sequences, which are one piece, into small amino acid groups, that is, words (k-mers). In addition, a preprocessing process that will provide high success is carried out with vector assignments and sequence digitization to be made according to this proximity information. When looking at the literature, the lengths of these amino acid groups are usually chosen in the range of 3 to 6 characters [40], [41].

After the sequences are divided into k-mers, it is necessary to analyze the closeness of these k-mers with each other. To avoid any k-mers combinations on all sequences and reveal different affinities, each sequence should be divided into k-mers 3 times by shifting one character from the first character 3 times. In this way, many k-mers are obtained, just as if 3 different sequences were processed [23]. In addition, the relationship between amino acids is confirmed 3 times. An example sequence-k-mer conversion is given in Figure 5.

EXAMPLE SEQUENCE

MDTSRPGAFVLSSAPLAALHNMAEMKT...

WORDS (K-MERS)

MDT, SRP, GAF, VLS, SAP, LAA, LHN, MAE, ...
DTS, RPG, AFV, LSS, APL, AAL, HNM, AEM, ...
TSR, PGA, FVL, SSA, PLA, ALH, NMA, EMK, ...

Fig. 5. An example sequence-k-mer conversion.

If sequences are split into words with more characters (e.g., 5-mer, etc.) instead of 3-mers, there will be fewer word combinations as well as splits in the vector representations of motifs as fewer words will remain in each motif. These divisions will cause the motifs to be partially lost and the model's success to decrease.

In order to analyze the proximity of words in sequences divided into k-mers and digitize the sequences, vector assignments to these k-mers should be done with the Word2Vec model. Word2Vec is an unsupervised neural network model that represents words in vector space. According to the prepared vocabulary, the words in the sequences are given points according to their similarity, and their probability of coming before or after each other is determined. Word2Vec has two different architectures, Continuous Bag-of-Words (CBOW) and Continuous Skip-Gram; although both of these architectures do the same job, there is a fundamental difference between them. CBOW determines the word in the center according to the previous and next word as much as the specified window size. Skip-Gram, on the other hand, predicts the previous and following words by the window size according to the central word [42].

In the tests performed in this study, 6 was chosen for the CBOW architecture and window size for representation. The vector size in which each k-mer will be represented in the model has been determined as the default value of 300. Figure 6 shows the distribution of k-mer affinities in the vector space at window size 6 of the Word2Vec model.

Just as the resolution of all images is determined as the same in deep learning models working with images, the same length of sequences in models working with protein sequences in text structure is a factor that increases success. However, choosing a shorter than required length while bringing the sequences to a fixed length may cause partial or complete loss of protein motifs and, thus, classification errors. The selection made longer than necessary will cause a waste of resources and time, as well as classification errors, since the processing will be done with strings filled with too many characters. For this reason, it is necessary to analyze the data set well and choose the right

length. Figure 7 shows the distribution of sequences in the data set by length.

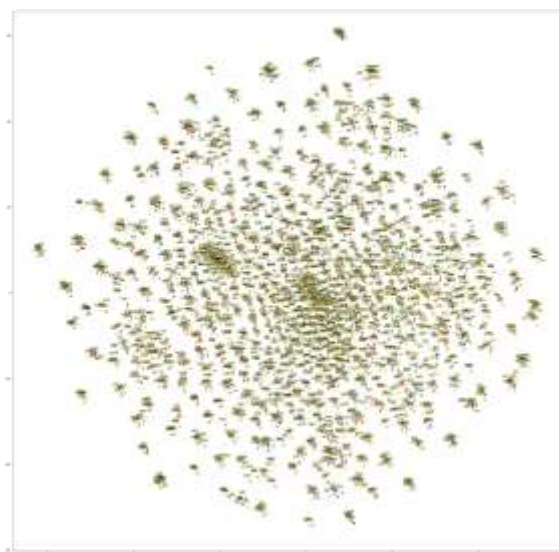


Fig. 6. Proximity representation of the Word2Vec model with 6 window sizes.

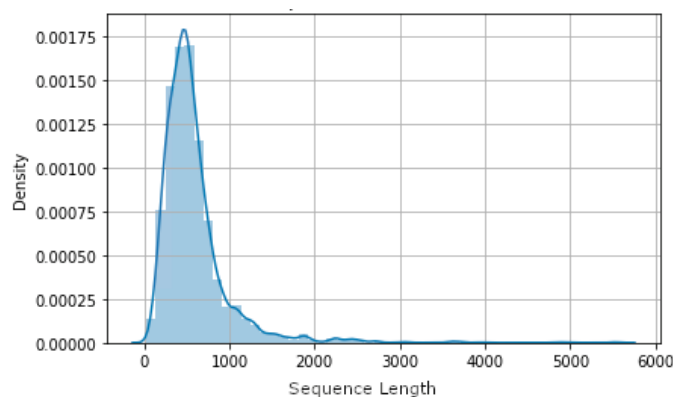


Fig. 7. Distribution of sequences in the data set according to their lengths.

Another purpose of separating sequences into words is to get shorter sequences from longer sequences. Looking at Figure 5, it can be seen that the majority of sequences are 400 to 600 in length. In this case, choosing the length of the arrays within this range will give the best results. Considering that the sequences are converted to 3-mers and digitized with vectors, it can be seen that synchronizing the sequences in their new form with 200 characters in length means equalizing the raw sequences in 600 characters. Sequences divided into K-mers are shortened to 200 words, long sequences are shortened from the end, and short sequences are extended to 200 words by adding the required number of 0s at the end.

The digitized and equal-length sequences are now ready to be given to the input embedding layer of the prepared deep learning model. In addition, the proximity information generated by training the Word2Vec model is also given to the "weights" parameter of the embedding layer so that the model starts training with these closeness values instead of random weights, which is aimed to perform a more successful training in a shorter time.

C. Long Short-Term Memory and Gated Recurrent Unit

LSTM is a new generation network developed against the problem that RNN forgets long-term information and dependencies [43]. As a solution to the continuous increase in gradients in RNN, a forget gate was added to each unit of the network, so that the network can decide which information to forget and which to keep [43], [44]. This structure uses tanh and sigmoid activation functions [45]. Similarly, GRU is a network designed to keep long and short-term dependencies in its memory. Although the internal structure of the units is very similar to the LSTM units, LSTM has input, output, and forget gates, while GRU has input and output gates. In GRU, the forget function is done with a forget key [46]. With all these functions, LSTM and GRU can decide which long- and short-term dependencies are remembered and which are forgotten by controlling the flow of information [47]. Although the GRU has almost the same level of success as the LSTM, it has been observed that the GRU mostly completes the training process faster than the LSTM because the GRU has 2 gates instead of 3 in the experiments.

D. Performance Metrics

The success of the models prepared in classification problems and other studies should be determined by various criteria. One of the most commonly used tools for determining success is the confusion matrix [48], and the accuracy, sensitivity, specificity, f-score, and average-precision values are calculated from this matrix [49]. In addition, these values will be used in drawing the Receiver Operating Characteristic (ROC) Curve [50]. While developing the models, the data sets are divided into parts in certain proportions and training, validation and testing steps are performed. Formulas for performance metrics are given in Equation 1-6.

$$\text{Accuracy (Acc)} = \frac{TP + TN}{TP + FP + FN + TN} \quad (1)$$

$$\text{Sensitivity (Se)} = \frac{TP}{TP + FN} \quad (2)$$

$$\text{Specificity (Sp)} = \frac{TN}{TN + FP} \quad (3)$$

$$\text{Precision (Pre)} = \frac{TP}{TP + FP} \quad (4)$$

$$F - \text{score} = \frac{2 * TP}{2 * TP + FP + FN} \quad (5)$$

$$\text{Average - Precision (AP)} = \frac{TP}{TP + TN} \quad (6)$$

While developing the models, the data sets are divided into parts in certain proportions, and training, validation, and testing steps are performed. In K-fold cross-validation, the data set is divided into k pieces, and k-1 pieces are used for training and 1 for testing at each step. In this way, one of the most accurate evaluations is performed as the models are trained and tested with different data many times [51], [52].

IV. RESULTS

All training and testing processes of the proposed model, the design details given in Table I, were carried out in Google Colab Pro with version 3.8 of the Python programming language. The primates TF protein dataset used in the training and testing of the model consists of 36243 protein sequences. Of these sequences, 70%, i.e., 25370, were used for training, 15% (5437) for validation, and 15% (5436) for testing.

Samples from all 72 classes are also available in all train, validation, and test datasets. This partitioning is done completely randomly via Python's Sklearn library. In order to achieve the highest success in training the model, an early stop function was added [38], and the fault tolerance was determined as 3 epochs. In order to determine the amount of data to be processed in unit time, the value of 256 was given to the batch_size parameter, taking into account the sequence lengths.

Two different models were studied for the effect of model layers and the number of units in each layer on model success and running time. The structures of these models are given in Table I.

TABLE I
BASIC STRUCTURES OF THE IMPLEMENTED MODELS

Model No	Model Structure
M1	Bidirectional LSTM (128), Bidirectional LSTM (256), Bidirectional GRU (256)
M2	Bidirectional LSTM (128), Bidirectional LSTM (128), Bidirectional LSTM (128), Bidirectional GRU (128)

The accuracy, precision, recall, f-score, and train time results of these two implemented models are given in Table II.

TABLE II
TEST RESULTS IMPLEMENTED MODELS

Model	Accuracy (%)	Precision (%)	Recall (%)	F-score (%)
M1	96.85	97.77	93.67	95.47
M2	97.96	97.55	95.26	96.22

When Table II is examined, it can be concluded that the model numbered M2 is a more successful classification. This result also indicates that models with more tiers and fewer units per tier perform better than models with fewer but more units per tier. The 5-fold cross-validation results of these two models are given in Table III.

TABLE III
5-FOLD CROSS VALIDATION RESULTS OF IMPLEMENTED MODELS

Model No	5-Fold Cross Validation Accuracy Results (%)
M1	96.67
M2	97.42

Table III also supports the result in Table II, showing that multilayer models with fewer units are more successful in this study. For this reason, in the continuation of the article, the model numbered M2 was determined as the proposed model. The reason for choosing 5-fold in this study is that the data set size is not very large.

In the proposed model M2, 3 layers of bidirectional LSTM and 1 layer of bidirectional GRU were used after the embedding layer, where sequences represented by vectors using Word2Vec and these vector affinity values were given as initial weights. 128 units were used in LSTM layers. A 0.3 dropout [53] layer has been added to prevent overfitting. The network of the next 128-unit bidirectional layer is determined as a bidirectional GRU for the model to complete a faster training process. To prevent overfitting, a dropout [53] layer of 0.3 was added, and then the flatten function was used. At the output of this layer, first, a dense layer with a value of 256, then a dropout [53] layer of 0.4 to prevent overfitting, and the model was finalized with

a dense layer of the class number to classification. The details of the model are shown in Table IV.

TABLE IV
DETAILED STRUCTURE OF THE PROPOSED MODEL

Layer (type)	Output Shape	Param #
Embedding	(None, 200, 300)	2540100
Bidirectional LSTM (128)	(None, 200, 256)	439296
Bidirectional LSTM (128)	(None, 200, 256)	394240
Bidirectional LSTM (128)	(None, 200, 256)	394240
Dropout (0.3)	(None, 200, 256)	0
Bidirectional GRU (128)	(None, 200, 256)	296448
Dropout (0.3)	(None, 200, 256)	0
Flatten	(None, 51200)	0
Dense (256)	(None, 256)	13107456
Dropout (0.4)	(None, 256)	0
Dense (72) (Classification)	(None, 72)	18504

The default learning rate value of 0.01 was used in the LSTM and GRU layers of the model. ReLU [54] was used as the activation function in the first dense layer, and Softmax [55] was used in the second because it was the classification layer. Since the data set is multi-class, Categorical Crossentropy [56] was used as the loss function, and Adam [57] was used as the optimizer.

The train and validation graphs for accuracy and loss of the model M1 are given in Figure 8, and the train and validation accuracy and loss graphs of the proposed model M2 are given in Figure 9.

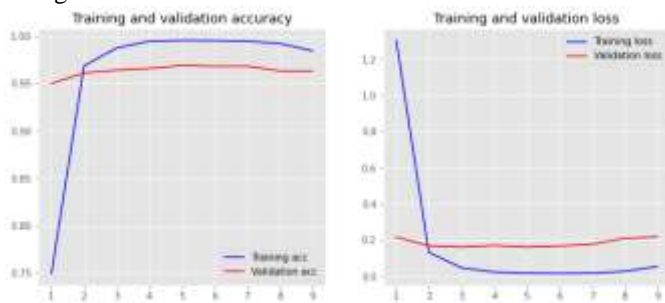


Fig. 8. Accuracy and loss graphs for train and validation of the model M1.

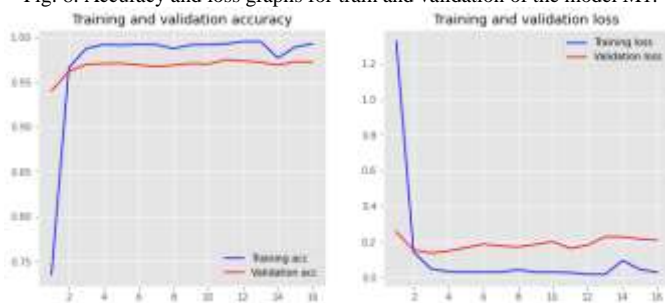


Fig. 9. Accuracy and loss graphs for train and validation of the proposed model M2.

The test results of the proposed model, the multi-layered structure of the model, which is prepared with layers with fewer units, increases the model's success and greatly shortens the training time. Figure 10 shows the ROC graph of the model M1. Figure 11 shows the ROC graph of the proposed model M2.

Some extension of Receiver operating characteristic to multi-class

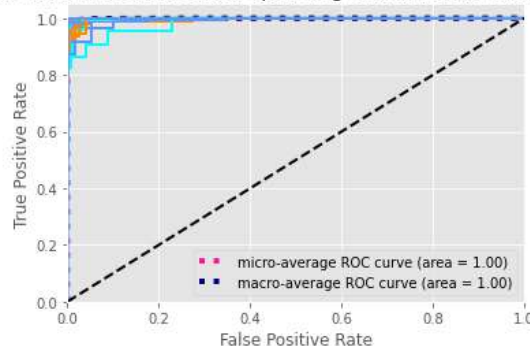


Fig. 10. ROC curve graph of the model M1.

Some extension of Receiver operating characteristic to multi-class

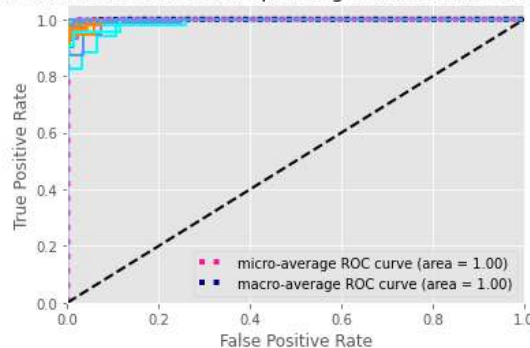


Fig. 11. ROC curve graph of the proposed model M2.

All the curves of the ROC graph in Figure 11 are gathered in the upper left corner, and their values are close to 1.0. This graph shows that the proposed model classifies 72 different classes with high success.

V. DISCUSSIONS

Along with the developments in computer science, computer science-supported studies have started to be carried out in bioinformatics. Statistical-based studies such as HMM can be among the early-mid-term examples of bioinformatics. Then, artificial neural networks and machine learning studies were used in bioinformatics. Today, various deep learning-based studies are used in bioinformatics and various biological and genetic data. In this study, transcription factor proteins in primates were classified by preparing a hybrid model combining the light and fast structure of RNN-based LSTM and GRU networks with the success of the Word2Vec model in preprocessing and similarity analysis. The data set was automatically compiled from public bioinformatics data and prepared originally. In this way, it was ensured that a deficiency in the literature was completed and an important and new contribution to the literature was made. In this process, a proximity analysis was performed by first dividing the sequences into small protein groups, namely k-mers. In this way, the model, which started its education ahead of similar models, achieved a more successful result in less epochs. Then, 3 LSTM layers and 1 GRU layer with fewer units per layer used were also used with lighter weight and more layers, and this design also increased the model's success while reducing the training time. In addition to all these contributions, the prepared

primates TF protein dataset was prepared and stored with the newly trained Word2Vec model and model weights for use in future studies.

Although there is no direct study on the classification of TF proteins in primates when the literature is scanned, the comparison with studies using similar data for the evaluation of the study will highlight the success of this study. Table V shows the comparison of the proposed model with some studies in the literature.

TABLE V
COMPARISON OF THE PROPOSED MODEL WITH VARIOUS STUDIES
IN THE LITERATURE

Authors	Dataset	Accuracy (%)
Le et al, 2019 [24]	Vesicular transport proteins	85.8
Belzen et al., 2019 [30]	CAFA3	93.7
Bileschi et al., 2022 [28]	Pfam Seed	95.8
Proposed model (M2)	Primates TF Proteins (original)	97.96

VI. CONCLUSION

Within the scope of this study, the classification of primates TF proteins, which has no precedent in the literature, was carried out. The model completed the classification process with an accuracy of 97.96% and achieved a high success rate in the literature. However, although plant TF proteins have been studied before in the literature [38], there are still deficiencies in the literature regarding the classification of TF proteins of other animals and organisms besides primates. In future studies, more successful classification of the TF proteins of primates classified in this study and the development of preprocessing steps will be provided, and it will be possible to study on the classification of TF proteins in other kingdoms.

REFERENCES

- [1] J. J. Shu, "A new integrated symmetrical table for genetic codes," *Biosystems*, vol. 151, pp. 21–26, Jan. 2017, doi: 10.1016/J.BIOSYSTEMS.2016.11.004.
- [2] J. D. WATSON and F. H. C. CRICK, "Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid," *Nature*, vol. 171, no. 4356, pp. 737–738, Apr. 1953, doi: 10.1038/171737a0.
- [3] D. R. Ferrier, "Protein Yapısı ve İşlevi," in *Lippincott Biyokimya: Görsel Anlatımlı Çalışma Kitapları*, B. A. Jameson, Ed. İstanbul: Nobel Tıp Kitapevleri, 2019, pp. 1–68.
- [4] Pfam, "Family: HLH (PF00010)." <http://pfam.xfam.org/family/pf00010> (accessed Feb. 02, 2019).
- [5] T. Kaplan and M. D. Biggin, "Quantitative Models of the Mechanisms that Control Genome-Wide Patterns of Animal Transcription Factor Binding," *Methods Cell Biol.*, vol. 110, pp. 263–283, Jan. 2012, doi: 10.1016/B978-0-12-388403-9.00011-4.
- [6] D. S. Latchman, "Transcription factors: an overview Function of transcription factors," *Int. J. Exp. Path.*, vol. 74, pp. 417–422, 1993.
- [7] M. Karin, "Too many transcription factors: positive and negative interactions," *New Biol.*, vol. 2, no. 2, pp. 126–131, 1990.
- [8] D. S. Latchman, "Transcription factors: An overview," *Int J Biochem Cell Biol*, vol. 29, no. 12, pp. 1305–1312, Dec. 1997, doi: 10.1016/S1357-2725(97)00085-X.
- [9] D. Petrey and B. Honig, "Is protein classification necessary? Toward alternative approaches to function annotation," *Curr Opin Struct Biol*, vol. 19, no. 3, pp. 363–368, Jun. 2009, doi: 10.1016/J.SBI.2009.02.001.
- [10] P. Baldi and S. Brunak, *Bioinformatics, Second Edition: The Machine Learning Approach*. Cambridge: MIT Press, 2001.
- [11] S. R. Eddy, "Hidden Markov models," *Curr Opin Struct Biol*, vol. 6, no. 3, pp. 361–365, Jun. 1996, doi: 10.1016/S0959-440X(96)80056-X.
- [12] M. M. Gromiha, "Protein Sequence Analysis," *Protein Bioinformatics*, pp. 29–62, Jan. 2010, doi: 10.1016/B978-8-1312-2297-3.50002-3.
- [13] S. F. Altschul, W. Gish, W. Miller, E. W. Myers, and D. J. Lipman, "Basic local alignment search tool," *J Mol Biol*, vol. 215, no. 3, pp. 403–410, Oct. 1990, doi: 10.1016/S0022-2836(05)80360-2.
- [14] M. N. Price *et al.*, "Mutant phenotypes for thousands of bacterial genes of unknown function," *Nature*, vol. 557, no. 7706, p. 503–509, May 2018, doi: 10.1038/s41586-018-0124-0.
- [15] N. Strodthoff, P. Wagner, M. Wenzel, and W. Samek, "UDSMProt: universal deep sequence models for protein classification," *Bioinformatics*, vol. 36, no. 8, pp. 2401–2409, Apr. 2020, doi: 10.1093/bioinformatics/btaa003.
- [16] K. S. Naveenkumar, B. R. Mohammed Harun, R. Vinayakumar, and K. P. Soman, "Protein Family Classification using Deep Learning," *bioRxiv*, p. 414128, Jan. 2018, doi: 10.1101/414128.
- [17] X. Du, Y. Cai, S. Wang, and L. Zhang, "Overview of deep learning," in *2016 31st Youth Academic Annual Conference of Chinese Association of Automation (YAC)*, 2016, pp. 159–164, doi: 10.1109/YAC.2016.7804882.
- [18] M. Huerta, F. Haseltine, Y. Liu, G. Downing, and B. Seto, "NIH working definition of bioinformatics and computational biology," Jul. 2000.
- [19] Q. Gong, W. Ning, and W. Tian, "GoFDR: A sequence alignment based method for predicting protein functions," *Methods*, vol. 93, pp. 3–14, Jan. 2016, doi: 10.1016/J.YMETH.2015.08.009.
- [20] H. bin Shen and K. C. Chou, "EzyPred: A top-down approach for predicting enzyme functional classes and subclasses," *Biochem Biophys Res Commun*, vol. 364, no. 1, pp. 53–59, Dec. 2007, doi: 10.1016/J.BBRC.2007.09.098.
- [21] A. Dalkiran, A. S. Rifaioglu, M. J. Martin, R. Cetin-Atalay, V. Atalay, and T. Doğan, "ECPred: a tool for the prediction of the enzymatic functions of protein sequences based on the EC nomenclature," *BMC Bioinformatics*, vol. 19, no. 1, p. 334, 2018, doi: 10.1186/s12859-018-2368-y.
- [22] D. Cozzetto, F. Minneci, H. Currant, and D. T. Jones, "FFPred 3: feature-based function prediction for all Gene Ontology domains," *Sci Rep*, vol. 6, no. 1, p. 31865, 2016, doi: 10.1038/srep31865.
- [23] E. Asgari and M. R. K. Mofrad, "Continuous Distributed Representation of Biological Sequences for Deep Proteomics and Genomics," *PLoS One*, vol. 10, no. 11, Nov. 2015.
- [24] N. Q. K. Le, E. K. Y. Yapp, N. Nagasundaram, M. C. H. Chua, and H. Y. Yeh, "Computational identification of vesicular transport proteins from sequences using deep gated recurrent units architecture," *Comput Struct Biotechnol J*, vol. 17, pp. 1245–1254, Jan. 2019, doi: 10.1016/J.CSBJ.2019.09.005.
- [25] F. G. Furat and T. Ibrikci, "Classification of Down Syndrome of Mice Protein Dataset on MongoDB Database," *Balkan Journal of Electrical and Computer Engineering*, pp. 44–49, Apr. 2018, doi: 10.17694/bajece.419553.
- [26] İ. ÖZER, "Classification of Urease Activity in Full-Fat Soybean Production by Extrusion Using Machine Learning Algorithms," *Balkan Journal of Electrical and Computer Engineering*, Aug. 2021, doi: 10.17694/bajece.941007.
- [27] S. Li, J. Chen, and B. Liu, "Protein remote homology detection based on bidirectional long short-term memory," *BMC Bioinformatics*, vol. 18, no. 1, p. 443, 2017, doi: 10.1186/s12859-017-1842-2.
- [28] M. L. Bileschi *et al.*, "Using deep learning to annotate the protein universe," *Nat Biotechnol*, vol. 40, no. 6, pp. 932–937, Jun. 2022, doi: 10.1038/s41587-021-01179-w.
- [29] R. Rao *et al.*, "Evaluating Protein Transfer Learning with TAPE," *Adv Neural Inf Process Syst*, vol. 32, pp. 9689–9701, Dec. 2019, [Online]. Available: <https://pubmed.ncbi.nlm.nih.gov/33390682>
- [30] J. Upmeyer zu Belzen *et al.*, "Leveraging implicit knowledge in neural networks for functional dissection and engineering of proteins," *Nat Mach Intell*, vol. 1, no. 5, pp. 225–235, 2019, doi: 10.1038/s42256-019-0049-9.
- [31] M. Torrisi, G. Pollastri, and Q. Le, "Deep learning methods in protein structure prediction," *Comput Struct Biotechnol J*, vol. 18, pp. 1301–1310, Jan. 2020, doi: 10.1016/j.csbj.2019.12.011.
- [32] S. Lim *et al.*, "A review on compound-protein interaction prediction methods: Data, format, representation and model," *Comput Struct Biotechnol J*, vol. 19, pp. 1541–1556, Jan. 2021, doi: 10.1016/J.CSBJ.2021.03.004.
- [33] C. Gustafsson, J. Minshall, S. Govindarajan, J. Ness, A. Villalobos, and M. Welch, "Engineering genes for predictable protein expression," *Protein Expr Purif*, vol. 83, no. 1, pp. 37–46, May 2012, doi: 10.1016/J.JEP.2012.02.013.

- [34] Pfam, "HSF-type DNA-binding PF00447." <https://www.ebi.ac.uk/interpro/entry/pfam/PF00447/logo/> (accessed Sep. 11, 2022).
- [35] H. Hu, Y.-R. Miao, L.-H. Jia, Q.-Y. Yu, Q. Zhang, and A.-Y. Guo, "AnimalTFDB 3.0: a comprehensive resource for annotation and prediction of animal transcription factors," *Nucleic Acids Res*, vol. 47, no. D1, pp. D33–D38, Jan. 2019, doi: 10.1093/nar/gky822.
- [36] IUPAC-IUB Comm. on Biochem. Nomenclature, "A one-letter notation for amino acid sequences. Tentative rules," *Biochemistry*, vol. 7, no. 8, pp. 2703–2705, Aug. 1968, doi: 10.1021/bi00848a001.
- [37] D. Ofer, N. Brandes, and M. Linial, "The language of proteins: NLP, machine learning & protein sequences," *Comput Struct Biotechnol J*, vol. 19, pp. 1750–1758, Jan. 2021, doi: 10.1016/J.CSBJ.2021.03.022.
- [38] A. B. Oncul, Y. Celik, N. M. Unel, and M. C. Baloglu, "Bhlhdb: A next generation database of basic helix loop helix transcription factors based on deep learning model," *J Bioinform Comput Biol*, Jun. 2022, doi: 10.1142/S0219720022500147.
- [39] B. Ay Karakuş, M. Talo, İ. R. Hallaç, and G. Aydin, "Evaluating deep learning models for sentiment classification," *Concurr Comput*, vol. 30, no. 21, pp. 1–14, Nov. 2018, doi: 10.1002/cpe.4783.
- [40] J. K. Vries, X. Liu, and I. Bahar, "The relationship between N-gram patterns and protein secondary structure," *Proteins: Structure, Function, and Bioinformatics*, vol. 68, no. 4, pp. 830–838, May 2007, doi: 10.1002/prot.21480.
- [41] J. K. Vries and X. Liu, "Subfamily specific conservation profiles for proteins based on n-gram patterns," *BMC Bioinformatics*, vol. 9, no. 1, p. 72, Dec. 2008, doi: 10.1186/1471-2105-9-72.
- [42] T. Mikolov, K. Chen, G. Corrado, and J. Dean, "Efficient Estimation of Word Representations in Vector Space," Jan. 2013.
- [43] K. Greff, R. K. Srivastava, J. Koutnik, B. R. Steunebrink, and J. Schmidhuber, "LSTM: A Search Space Odyssey," *IEEE Trans Neural Netw Learn Syst*, vol. 28, no. 10, pp. 2222–2232, Oct. 2017, doi: 10.1109/TNNLS.2016.2582924.
- [44] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, May 2015, doi: 10.1038/nature14539.
- [45] G. van Houdt, C. Mosquera, and G. Nápoles, "A review on the long short-term memory model," *Artif Intell Rev*, vol. 53, no. 8, pp. 5929–5955, Dec. 2020, doi: 10.1007/s10462-020-09838-1.
- [46] Y. Gao and D. Glowacka, "Deep Gate Recurrent Neural Network," in *Proceedings of The 8th Asian Conference on Machine Learning*, Jul. 2016, vol. 63, pp. 350–365. [Online]. Available: <https://proceedings.mlr.press/v63/gao30.html>
- [47] A. Şeker, B. Diri, and H. H. Balık, "Derin Öğrenme Yöntemleri ve Uygulamaları Hakkında Bir İnceleme," *Gazi Mühendislik Bilimleri Dergisi*, vol. 3, no. 3, pp. 47–64, Nov. 2017.
- [48] C. Sammut and G. I. Webb, Eds., *Encyclopedia of Machine Learning*. Boston, MA: Springer US, 2010. doi: 10.1007/978-0-387-30164-8.
- [49] A. Luque, A. Carrasco, A. Martín, and A. de las Heras, "The impact of class imbalance in classification performance metrics based on the binary confusion matrix," *Pattern Recognit*, vol. 91, pp. 216–231, Jul. 2019, doi: 10.1016/J.PATCOG.2019.02.023.
- [50] B. Ozenne, F. Subtil, and D. Maucourt-Boulch, "The precision–recall curve overcame the optimism of the receiver operating characteristic curve in rare diseases," *J Clin Epidemiol*, vol. 68, no. 8, pp. 855–859, Aug. 2015, doi: 10.1016/J.JCLINEPI.2015.02.010.
- [51] A. Rohani, M. Taki, and M. Abdollahpour, "A novel soft computing model (Gaussian process regression with K-fold cross validation) for daily and monthly solar radiation forecasting (Part: I)," *Renew Energy*, vol. 115, pp. 411–422, Jan. 2018, doi: 10.1016/j.renene.2017.08.061.
- [52] Z. Xiong, Y. Cui, Z. Liu, Y. Zhao, M. Hu, and J. Hu, "Evaluating explorative prediction power of machine learning algorithms for materials discovery using k-fold forward cross-validation," *Comput Mater Sci*, vol. 171, p. 109203, Jan. 2020, doi: 10.1016/j.commatsci.2019.109203.
- [53] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A Simple Way to Prevent Neural Networks from Overfitting," *Journal of Machine Learning Research*, vol. 15, no. 56, pp. 1929–1958, 2014, [Online]. Available: <http://jmlr.org/papers/v15/srivastava14a.html>
- [54] L. Parisi, D. Neagu, R. Ma, and F. Campean, "Quantum ReLU activation for Convolutional Neural Networks to improve diagnosis of Parkinson's disease and COVID-19," *Expert Syst Appl*, vol. 187, p. 115892, Jan. 2022, doi: 10.1016/j.eswa.2021.115892.
- [55] A. Basturk, M. E. Yuksei, H. Badem, and A. Caliskan, "Deep neural network based diagnosis system for melanoma skin cancer," in *2017 25th Signal Processing and Communications Applications Conference (SIU)*, May 2017, pp. 1–4. doi: 10.1109/SIU.2017.7960563.
- [56] 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, Aug. 2018, doi: 10.1007/s13244-018-0639-9.
- [57] E. YAZAN and M. F. Talu, "Comparison of the stochastic gradient descent based optimization techniques," in *2017 International Artificial Intelligence and Data Processing Symposium (IDAP)*, Sep. 2017, pp. 1–5. doi: 10.1109/IDAP.2017.8090299.

BIOGRAPHIES



ALİ BURAK ÖNCÜL Amasya, in 1991. He received the B.S. in 2013 and M.S. degrees in 2016 in computer engineering from Ondokuz Mayıs University, Samsun, and the Ph.D. degree in computer engineering from Karabük University, Karabük, in 2022.

Since 2015, he has been a Research Assistant with the Computer Engineering Department, Kastamonu University. His areas of study are GNU/Linux, operating systems, deep learning, bioinformatics, protein structure and classifications.

A Software Realization of Disturbance Rejection Optimal FOPID Controller Design Methodology by Using Soft Computing Techniques

Sevilay Tufenkci, Baris Baykant Alagoz, Celaledin Yeroglu, Bilal Senol

Abstract— This study presents a soft computing tool for the computer-aided design of disturbance rejection FOPID controllers based on the maximization of Reference to Disturbance Ratio (RDR) index. The study illustrates the utilization of software routines to implement a soft computing scheme in order to solve a closed loop disturbance rejection FOPID control system design problem for a target gain margin specification. Authors demonstrate that the complex design efforts, which involve a high level of mathematical knowledge, can be easily performed by using basic software routines when soft computing techniques are employed effectively in the computation processes. Illustrative design examples are shown to show effectiveness of the proposed design method.

Index Terms— Disturbance rejection control, FOPID controller, genetic algorithm, phase margin, RDR

I. INTRODUCTION


FRACTIONAL CALCULUS has come out at the end of the 17th century, and today it becomes a popular mathematical tool to solve modeling and design problems in engineering and applied science [1-3]. However, its utilization in control engineering practice has been delayed due to the

computational load of fractional order operators. Approximate realization of fractional order operators and development of soft computation tools contributes to the solution of these computationally complex design problems.


Time-domain design methodologies can yield more relevant results for the control system design practice because they perform a design effort based on the time response of the control systems. This design strategy is practically reasonable because the control system performances of the real systems are generally evaluated according to their time responses by using test signals that are applied to the system. However, derivation of the analytical controller design rules regarding the time responses of higher order control systems or fractional order systems is rather difficult to obtain in the time domain design, and this turns into an important design limitation for development of analytical tuning rules in the time domain for complex system models. However, frequency domain analysis has an important role in control theory, since complex control systems can be easily studied in the frequency domain [4]. In the frequency domain, there are several graphical methods that can work not only with low-order systems but also with high-order systems and/or fractional order systems. The main reasons for preferring the frequency domain in the analysis and design of fractional order control systems are that stability analysis and analytical solutions in order to obtain controller design rules are more straightforward in the frequency domain due to allowing rather simple expressions of the linear time invariant high-order system models in the frequency domain compared to the time domain models. For these reasons, intensive studies have been carried out especially on the controller design based on the frequency domain modeling [5]. To evaluate behavior of the control system in the frequency domain, several graphical tools such as the Bode diagram, the Nichols and Nyquist diagrams and the root locus plots have been used.

In the frequency domain controller design, the phase margin is an important measure in order to express the robustness of the system. The phase margin is also related to controller performance; for instance, it affects the damping rate of the


SEVILAY TUFENKCI, is with Department of Computer Technology Department, Malatya Turgut Ozal University, Malatya, Turkey (e-mail: sevilay.tufenkci@gmail.com)

 <https://orcid.org/0000-0001-9815-7724>


BARIS BAYKANT ALAGOZ, is with Department of Computer Engineering, Inonu University, Malatya, Turkey (e-mail: baris.alagoz@inonu.edu.tr).

 <https://orcid.org/0000-0001-5238-6433>

CELALEDDIN YEROGLU, is with Department of Computer Engineering, Inonu University, Malatya, Turkey (e-mail: c.yeroglu@inonu.edu.tr).

 <https://orcid.org/0000-0002-6106-2374>

BILAL SENOL, is with Department of Computer Engineering, Inonu University, Malatya, Turkey (e-mail: bilal.senol@inonu.edu.tr).

 <https://orcid.org/0000-0002-3734-8807>

Manuscript received Mar 27, 2022; accepted Nov 19, 2022.
DOI: [10.17694/bajece.1092971](https://doi.org/10.17694/bajece.1092971)

system [6]. Several controller design topics cover the phase margin together with the gain margin [6, 7]. The results associated with phase and gain margin can be graphically displayed by using the Bode diagrams.

In real-world control applications, the control systems are exposed to a number of disturbance effects due to the system noise or external factors acting on the control system. Those disturbances may be unpredictable and they may influence the control systems in real control actions and prevent them from showing the expected control performance in the real-world systems. Therefore, it is very substantial to minimize the impacts of disturbances on the system response while designing a controller for the real-world control applications. Various studies have been presented in order to reduce the undesirable effects of disturbances on the control performance and improve the robust stability of the control systems [8-11, 12, 13].

In this study, the reference to disturbance ratio (RDR) in cooperation with the gain margin specification is employed in the design task to increase the disturbance rejection capacity of closed loop control systems. The RDR measure is used to calculate the quantitative evaluation of the dominance of reference input signal on the input disturbance signal at the system output [14, 15, 27]. While a control system shows satisfactory disturbance rejection performance in the case of $RDR \gg 1$, the control system does not show any disturbance rejection skills and performance of the control system is vulnerable to the impacts of disturbances in the case of $RDR \leq 1$.

Today, closed loop control systems are widely used in many industrial control applications. The classical controller design process involves determining optimal controller coefficients in a way that responses of the designed control system meet several design specifications such as robust stability, disturbance rejection, fault tolerance etc. Therefore, multi-objective optimization methods have been widely used in the controller design problems. The optimization is the process of choosing the best possible tuning options under certain criteria. Many different optimization algorithms have been implemented to perform this task. The selection of a suitable optimization method is important in the design stage and assets of optimization methods have a role in optimal controller tuning problems. While one algorithm can find a satisfactory solution for the problem, the other optimization algorithm may not reach the desired result. Metaheuristic algorithms have been commonly utilized for obtaining the near-optimum solutions. They are preferred to reach the optimal solution in an acceptable time, particularly in the case of large-scale, mathematically not-well structured, complicated optimization problems. Today, metaheuristic optimization algorithms can be classified into several categories, for example the biology-based, the physics-based, swarm-based, the social-based etc. This algorithm has been frequently used in optimal tuning of fractional-order controls [16-18]. Recent works have been implemented the chaotic yellow saddle goatfish algorithm[19], hybrid Lévy flight

distribution and simulated annealing algorithm[20], a multi-objective genetic algorithm (MOGA) and particle swarm optimization[21]. These works reveal that metaheuristic methods can deal with the complicated optimization problems for controller design, However they haven't clearly stated and discussed the soft computation details in their works. In the current study, we present soft computation details that were used for the proposed design scheme.

Mathematical complication in the solution of constrained optimization problems can be alleviated by using soft constraints [22]. The soft constraints refer to a constraint that can be violated in the optimization problem at a penalty cost and the optimization algorithm can progressively continue for the reduction or removal of the penalty costs [22]. Hard constraints cannot be violated at any cost. For further theoretical consideration, a deepened theoretical discussion on the soft and hard constraints was presented in [23] and their roles in the optimization process [24]. However, the practice of the soft constraints has appeared and developed through the soft computing applications [25, 26].

Soft computation methods can facilitate very complicated design tasks in engineering problems. In this perspective, the main motivation of this study is the proposal and implementation of some soft computation techniques in an optimal disturbance rejection FOPID controller design problem. Analytical solution of this optimization problem is quite complicated because of difficulties in fractional calculus and the mathematical modeling of environmental disturbances. Authors aim to demonstrate that such complicated design tasks can be easily performed by using basic software routines. Different from the similar works, the study introduces these soft computation routines (soft constraints, crossover frequency calculation via zero crossing detection etc.) in detail, and illustrates an application of them in a disturbance reject FOPID design problem.

In the current study, fractional order control coefficients were adjusted by combining the phase margin constraint and RDR index. The RDR index is maximized for a gain margin specification in order to improve the disturbance rejection performance of the FOPID control systems. The phase margin specification is defined as the soft constraints of the optimization problem. Thus enables software realization of the complex phase margin analyses and eliminates a need for complicated mathematical derivations. Such a softening of the hard constraints also modifies the search spaces and contributes to search performance of the metaheuristic optimization methods by allowing progressive approximation to optimal solutions. For the solution of this optimal tuning problem, authors implemented the most fundamental and popular metaheuristic search method that is a genetic algorithm. Two illustrative FOPID controller design examples were shown to evaluate the efficiency of the proposed optimal disturbance reject FOPID controller design methodology.

II. PRELIMINARIES AND PROBLEM STATEMENTS

A. RDR Analysis for Disturbance Rejection

RDR analyses were suggested for quantitative evaluation of disturbance rejection capacity of the closed-loop control systems [14, 15, 27]. Consideration of the RDR index in fractional-order FOPID control designs allows improving disturbance rejection control performance [15, 27-31]. Effective disturbance rejection controller design approaches based on the RDR index have also been presented and their improvements for disturbance rejection control were shown [27-31].

The communication channel based analysis on the negative feedback loop yields the following RDR formulation for the control systems [14, 15, 29].

$$RDR(\omega) = |C(j\omega)|^2 \quad (1)$$

Since the RDR index can take very high values, their values are expressed in decibel(dB) as follows:

$$RDR_{dB}(\omega) = 20 \log |C(j\omega)| \quad (2)$$

To improve disturbance rejection performance, the RDR constraint that expresses a lower boundary in the RDR spectrum was suggested for the operating frequency range $[\omega_{\min}, \omega_{\max}]$ of control systems [27] and a minimum RDR constraint was expressed as

$$\min\{RDR_{dB}(\omega)\} \geq M \text{ for } \omega \in [\omega_{\min}, \omega_{\max}]. \quad (3)$$

Figure 1 shows a closed loop FOPID control system with the additive input disturbance model $D(s)$ for the RDR index. The function $C(s)$ is the transfer function of FOPID controller and it is widely expressed as

$$C(s) = k_p + k_d s^\mu + \frac{k_i}{s^\lambda}. \quad (4)$$

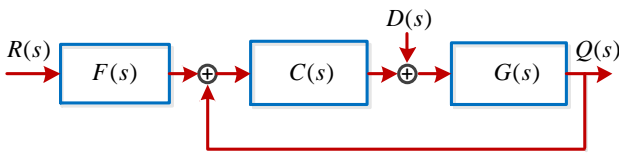


Fig.1. FOPID control system with ste-point filter

In contrast to the classical PID controller, a FOPID controller have two additional fractional order coefficients, which are fractional integrator orders λ and fractional derivative orders μ in addition to the k_p , k_d and k_i gain coefficients. These additional fractional order integral and derivative orders provide an opportunity to improve control system performance [3, 5]. By considering Equation (1), the RDR formulas of the FOPID control systems were derived as follows [15]

$$RDR(\omega) = (k_p + k_i \cos(\frac{\pi}{2}\lambda)\omega^{-\lambda} + k_d \cos(\frac{\pi}{2}\mu)\omega^\mu)^2 + (k_d \sin(\frac{\pi}{2}\mu)\omega^\mu - k_i \sin(\frac{\pi}{2}\lambda)\omega^{-\lambda})^2 \quad (5)$$

The phase margin is a well-known property in classical control systems and it is expressed as an important criterion for ensuring the robustness of the system stability [6, 7]. The phase margin also affects the closed-loop damping ratio of the second order systems [6].

The phase margin is widely used for the stabilization of control systems, and the general form of the phase margin (φ_m) for the open loop transfer function $L(s) = C(s)G(s)$ is written by

$$\varphi_m = \text{Arg}[G(j\omega_c)C(j\omega_c)] + \pi. \quad (6)$$

Here, ω_c refers to the crossover frequency that is defined as the angular frequency that satisfy $|G(j\omega_c)C(j\omega_c)| = 1$.

Therefore, the calculation of the phase margin requires a solution of nonlinear, complex valued equations and it causes difficulties in the mathematical solution of this problem. The phase margin basically represents the difference between the phase at the crossover frequency of the open loop control system and the $-\pi$ angle and rather easy to solve graphically.

B. A Brief Review of Genetic Algorithm

Genetic algorithm (GA) is one of the most popular metaheuristic search techniques, which can provide easy solutions for today's complex design problems. The GA is based on the natural selection principles and the algorithm was suggested by John Holland in 1975. The basic idea in this method is based on the survival of the good individuals (solutions) in the genetic pool (solution population). It tries to find the best result or the closest one to be the best in its search space [32-34]. The search mechanism of the genetic algorithm resembles the transmission of the physical and biological characteristics of living things to the next generation through the genes. As each generation is formed by the combination of better features from the previous generations, the best individual survives with higher chance in selection mechanisms and each generation gets better as generations progress. Fitness values of each individual are used in a selection mechanism that tends to select the better fitting individuals. After selection of more fitting individuals to the solution, the next generation of individuals is reproduced through a series of genetic processes such as mutation and crossover. The basic algorithmic steps are shown in Figure 2 [35]. The GA algorithm has become a very effective evolutionary computational method.

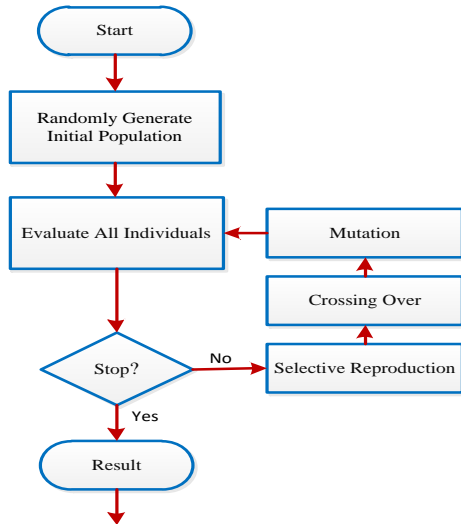


Fig.2. A fundamental flow chart for the genetic algorithm

III. SOFTWARE REALIZATION OF OPTIMAL FRACTIONAL ORDER PID CONTROL SYSTEM DESIGN

This section summarizes the mathematical background and the software realization of the computer-aided design method that are implemented for the optimal tuning problem of FOPID controls coefficients based on the maximization of RDR index and complying with the phase margin specification.

We consider a fractional plant function that is expressed in a general form as follows:

$$G(s) = \frac{a_0}{b_2 s^k + b_1 s^\alpha + b_0} \quad (7)$$

The closed loop transfer function, which is composed of the controlled system model (Equation (7)) and the fractional order PID controller (Equation (4)), can be written by

$$T(s) = \frac{Q(s)}{R(s)} = \frac{a_0 k_d s^{(\lambda+\mu)} + a_0 k_p s^\lambda + a_0 k_i}{b_2 s^{(\lambda+k)} + a_0 k_d s^{(\lambda+\mu)} + b_1 s^{(\lambda+\alpha)} + (b_0 + a_0 k_p) s^\lambda + a_0 k_i} \quad (8)$$

This transfer function is implemented by using the *fotf()* function of *fotf toolbox* [36]. Figure 3 shows the pseudocode that describes a software implementation of the closed loop FOPID control system model.

A pseudocode for RDR index calculation according to the equation (5) is shown in Figure 4. The code returns the minimum RDR value in a frequency range $([\omega_{\min}, \omega_{\max}])$ to implement

$$RDR_c(x) = \min_{RDR} \{ RDR_{dB}(\omega, x), \omega \in [\omega_{\min}, \omega_{\max}] \}.$$

The crossover frequency (ω_c) is commonly found by solving $|G(j\omega_c)C(j\omega_c)| = 1$. One can write this equation by forming open loop transfer function $L(s) = C(s)G(s)|_{s \rightarrow j\omega}$ in the frequency region and writing the magnitude of the resulting complex rational function by

$$\left| \frac{a_0 k_d (j\omega_c)^{\mu+\lambda} + a_0 k_p (j\omega_c)^\lambda + a_0 k_i}{b_2 (j\omega_c)^{k+\lambda} + b_1 (j\omega_c)^{\alpha+\lambda} + b_0 (j\omega_c)^\lambda} \right| = 1 \quad (9)$$

Algorithm *fractional order transfer function* is

Inputs: Coefficients of $G(s)$ a_0, b_2, b_1, b_0 ,

Order of $G(s)$ k, α

Coefficients of $C(s)$ k_p, k_d, k_i ,

Order of $C(s)$ u, λ

Coefficients of $T(s)$ $c_0, c_1, c_2, c_3, c_4, d_2, d_1, d_0$

Order of $T(s)$ $\beta_0, \beta_1, \beta_2, \beta_3$,

$\beta_4, \gamma_0, \gamma_1, \gamma_2$,

Outputs: Symbolic model of $T(s)$ T_s ,

% Coefficients of closed loop transfer function

$c_0 = a_0 * k_i$;

$c_1 = a_0 * k_p + b_0$;

$c_2 = a_0 * k_d$;

$c_3 = b_1$;

$c_4 = b_2$;

$\beta_0 = 0$;

$\beta_1 = \lambda$;

$\beta_2 = \lambda + u$;

$\beta_3 = \lambda + \alpha$;

$\beta_4 = k + \lambda$;

$d_2 = a_0 * k_d$;

$d_1 = k_p * a_0$;

$d_0 = k_i * a_0$;

$\gamma_2 = (\lambda + u)$;

$\gamma_1 = \lambda$;

$\gamma_0 = 0$;

% Software realization of the closed loop transfer function by using *fotf toolbox*

$T_s = \text{fotf}([c_4 \ c_3 \ c_2 \ c_1 \ c_0], [\beta_4 \ \beta_3 \ \beta_2 \ \beta_1 \ \beta_0], [d_2 \ d_1 \ d_0], [\gamma_2 \ \gamma_1 \ \gamma_0])$;

Return T_s

Fig.3. A pseudocode for implementation of fractional order transfer functions by using *fotf toolbox*

Algorithm *minimum RDR* is

Inputs: Coefficients of $C(s)$ k_p, k_d, k_i ,

Order of $C(s)$ u, λ ,

Angular frequency vector w ,

RDR value vector RDR_{dB}

Outputs: Minimum RDR value $\min RDR_{dB}$

% Calculation of logarithmic RDR values according equations(5)

$RDR_{dB} = 20 * \log_{10}((k_p + k_i * \cos((\pi/2) * \lambda) * w.^{(-\lambda)} + k_d * \cos((\pi/2) * u) * w.^{(u)} + (k_d * \sin((\pi/2) * u) * w.^{(u)} - k_i * \sin((\pi/2) * \lambda) * w.^{(-\lambda)}))^2)$;

% Calculation of minimum RDR index

$\min RDR_{dB} = \min(RDR_{dB})$;

Return $\min RDR_{dB}$

Fig.4. A pseudocode for RDR the index and the minimum RDR calculation for the realization of the equation (9)

The equation 9 requires a nonlinear equation solving method to find crossover frequency (ω_c). To solve this complex equation by using the soft computing technique, one can form a zero crossing by taking the logarithm of both side of the equation and obtain $f(\omega_c) = \log_{10} |G(j\omega_c)C(j\omega_c)| = 0$. To solve this equation and roughly find the sampled crossover frequency ω_c and the phase at the crossover frequency ($Arg[G(j\omega_i)C(j\omega_i)]$), a zero-crossing detection mechanism as depicted in Figure 5 is used. Thus, this nonlinear equation can be easily solved by using zero crossing detection with an error (tolerance) less than the unit sampling distance ($|\omega_i - \omega_{i-1}|$).

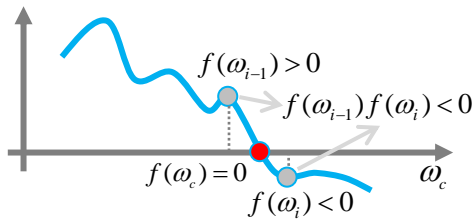


Fig.5. A soft computing technique for searching the crossover frequency ω_c via the zero crossing detection

Figure (6) shows a pseudocode for the zero-crossing detection algorithm that can be used for approximate calculation of the crossover frequency as in Figure 5.

Algorithm *zero-crossing detection* is
 Inputs: Value vector of the function f ,
 Angular frequency vector w ,
 Outputs: Crossover frequency wc

```

For i is from 1 to length of f
if f(i-1)*f(i)<0
wc=w(i);
end
Return wc

```

Fig.6. A pseudocode for the zero-crossing detection algorithm that can be used for approximate calculation of the crossover frequency

Then, the objective function to minimize is written as [29]

$$E(x) = \left(\frac{1}{RDR_c(x) + \varepsilon_0} \right)^2, \quad (10)$$

where $\varepsilon_0 > 0$ is a very small real number in order to avoid zero divisions in the case of a zero value of RDR_c . Here, the vector $x = [k_p, k_i, k_d, \lambda, \mu]$ is the controller coefficients. The standard definition of the optimization problem is written as

$$\min_x E(x), \quad (11)$$

$$\text{S.t.: } |\varphi_m - \varphi_T| < \varepsilon. \quad (12)$$

This objective function maximizes the RDR_c index that allows an improved input disturbance rejection performance for the additive input disturbance signal $D(s)$ (See Figure 1). The phase margin specification of the system is introduced by the inequality constraint. This phase margin constraint is necessary for the stabilization of the control system. The target phase margin value is set $\varphi_T = \frac{2\pi}{3}$ for robust stabilization of the system. According to Equation (12), the parameter φ_m stands for the phase margin of the current system, the φ_T denotes the target phase value. The ε is a small positive number to define an approximation of system phase margin (φ_m) to the target phase margin (φ_T) within the ε tolerance range. This range allows the enhancement of RDR performance during the optimization process of genetic algorithms in the expanse of a small allowable deviation from target phase margin. Thus it softens the phase constraint and facilitates the constrained optimization problem ((Equations (11) and (12))). Figure 7 shows a pseudocode to implement this optimization problem according to a soft constraint $|\varphi_m - \varphi_T| < \varepsilon$. In this code, when the phase constraint $|\varphi_m - \varphi_T| < \varepsilon$ is not satisfied, the absolute phase margin error $|\varphi_m - \varphi_T|$ is amplified by a factor of 10^{10} in order to reduce search possibility of such candidate solutions in the search space. Assignment of very high error values for these candidate solutions prevents the survival of them in next generations and eliminates these solutions that violate the soft constraint.

Algorithm *objective function* is
 Inputs: Target phase margin ϕ_m ,
 Calculated phase margin ϕ_t ,
 Minimum RDR value $\min RDR_{dB}$
 Outputs: Objective function to minimize E

```

if |phi_m- phi_t| < 0.01
E=(1/(( minRDRdB)^2)+0.001)
else
E=|phi_m- phi_t| *10^10
end
Return E

```

Fig.7. A pseudocode for solving optimization problem according to a soft constraint ($\varepsilon = 0.01, \varepsilon_0 = 0.001$)

A software realization of this design problem in Matlab is carried out by employing soft computing techniques. This avoids the requirement of solving very difficult analytical equations to develop practical computer-aided design tools. The Matlab codes that were written for this design tool are presented in the Appendix section.

IV. ILLUSTRATIVE EXAMPLES

In this section, two illustrative design examples are presented.

Example 1: Let's design a disturbance rejection FOPID control for the fractional order plant model given below [37]:

$$G(s) = \frac{1}{0.8s^{2.2} + 0.5s^{0.9} + 1} \quad (13)$$

The closed loop transfer function of the system is written as

$$T(s) = \frac{k_d s^{(\lambda+\mu)} + k_p s^\lambda + k_i}{0.8s^{(\lambda+2.2)} + k_d s^{(\lambda+\mu)} + 0.5s^{(\lambda+0.9)} + (1+k_p)s^\lambda + k_i} \quad (14)$$

Target phase margin for the control system is set as $\varphi_T = \frac{2\pi}{3}$. Optimal FOPID controller coefficients were

searched by using GA optimization for the fractional order controller parameter ranges LB = [1 1 1 0.3 0.3] and UB = [500 500 500 2 2] according to $x = [k_p, k_i, k_d, \lambda, \mu]$. The optimal disturbance reject FOPID controller was obtained as

$$C(s) = 307.4346 + 145.9923s^{0.8984} + \frac{452.6156}{s^{1.2533}} \quad (15)$$

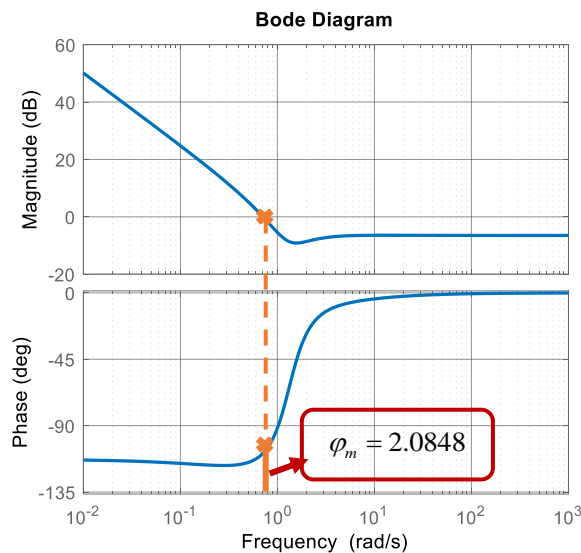


Fig.8. Bode diagram of the designed control system

The stability of systems can be evaluated by examining the Bode diagram. When the amplitude graph is 0 dB, the distance of the system phase to -180 degrees ($-\pi$ radian) represents the phase margin. In Figure 8, the phase margin is shown with a thick orange solid line. The phase margin value for the designed system is obtained as $\varphi_m = 2.0848$. Figure 9 shows the step responses of the proposed FOPID controller and Zhao et al.'s optimal FOPID controller. Figure 10 shows a close view of Figure 9 in order to compare step performances in terms of the rise time, the settling time and the maximum overshoot. The figure indicates improvement of the step response performance by means of the proposed FOPID controller in this example.

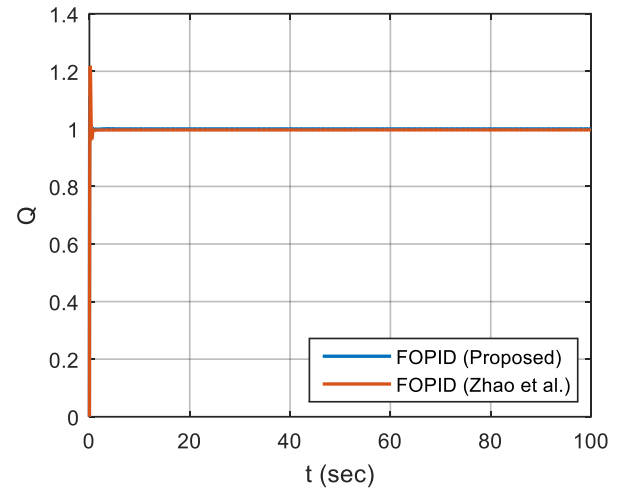


Fig.9. Step responses of the control systems

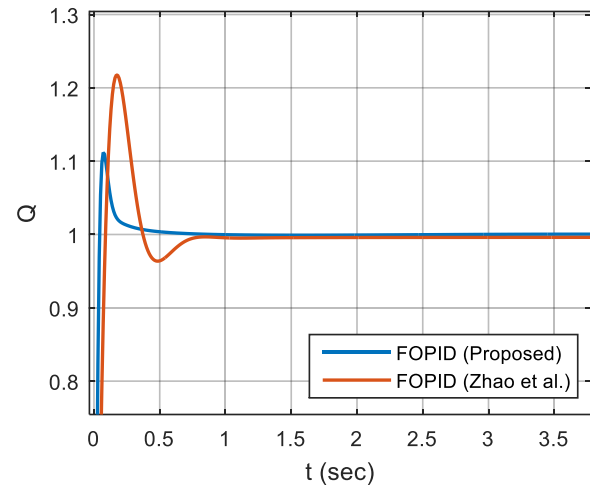


Fig.10. A close view of the step responses in Figure 9

Figure 11 shows the RDR spectrum of the FOPID control systems. Higher RDR values indicate the improved disturbance rejection control performance in the corresponding frequency. The proposed FOPID controller can provide higher RDR values in the majority of the RDR spectrum and this indicates the proposed FOPID controller can present a better disturbance rejection performance, particularly at the higher frequency region compared to Zhao et al.'s FOPID controller. To further improve setpoint control performance, authors used the 2DOF setpoint filter FOPID control system (See Figure 1) in disturbance rejection control simulations. The setpoint filter is configured as $F(s) = \frac{1}{3s+1}$

that avoids high overshoot step response when settling to the setpoint 1 as in Figure 12. (Zhao et al.'s FOPID control system does not include a set-point filter) These simulations were performed in the Simulink environment by using the *fotf* toolbox [36]. Figure 12 shows step disturbance responses of the controllers. An additive step disturbance was applied to the control system at 40 sec simulation time. Figure 13 shows a close view of disturbance rejection control for both controllers.

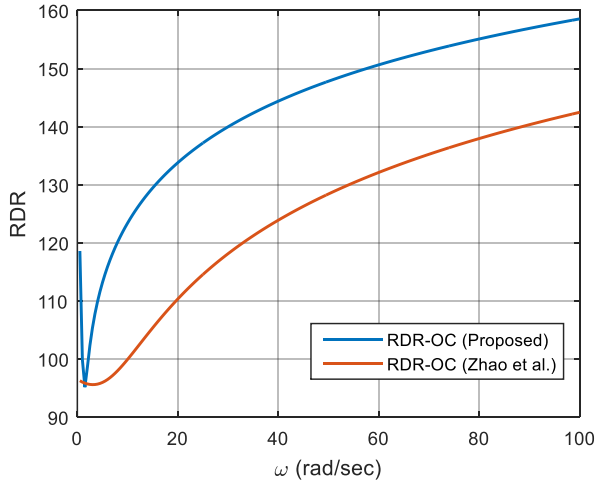


Fig.11. The RDR spectrums of the control systems

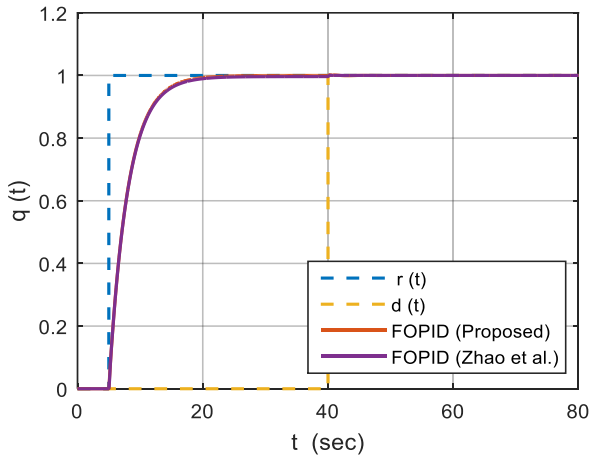


Fig.12. Step responses and disturbance responses of the FOPID controllers in case of an additive step disturbance insertion into the plant input at 40 sec

Figure 14 shows system responses when a sinusoidal disturbance signal with the amplitude of 1 and a frequency of 3.14 rad/sec was inserted into the control system. The figure shows that the disturbance rejection control performance is improved by the proposed FOPID control system in this frequency.

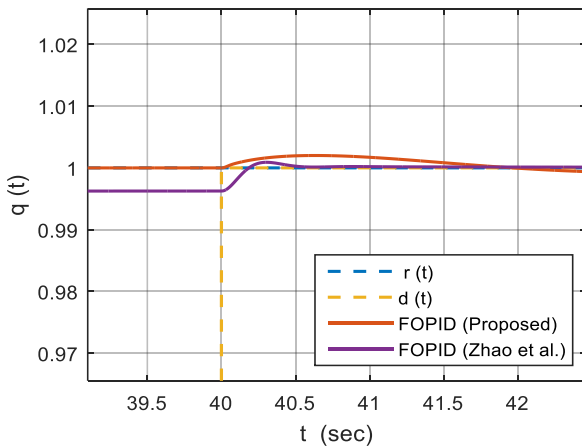


Fig.13. A close view of the disturbance responses in Figure 12

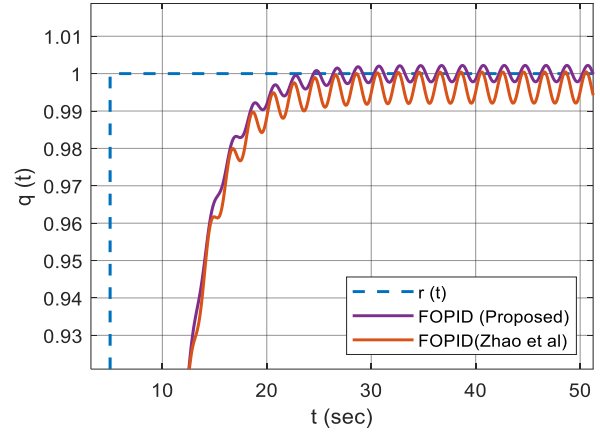


Fig.14. Sinusoidal disturbance responses of the control systems

Example 2: Let's design a disturbance rejection FOPID control for the fractional order plant model given below [38]:

$$G(s) = \frac{0.5}{1.5s^{1.3} + 1} \quad (16)$$

The closed loop transfer function of the system is written as

$$T(s) = \frac{0.5k_d s^{(\lambda+\mu)} + 0.5k_p s^\lambda + 0.5k_i}{0.5k_d s^{(\lambda+\mu)} + 1.5s^{(\lambda+1.3)} + (1+0.5k_p)s^\lambda + 0.5k_i} \quad (17)$$

A target phase margin to be achieved by using the control system is set as $\varphi_T = \frac{2\pi}{3}$. The optimal FOPID controller

coefficients obtained by using GA optimization for a FOPID parameter search ranges LB = [1 1 1 0.5 0.3] and UB = [50 50 50 2 2] according to $x = [k_p, k_i, k_d, \lambda, \mu]$. Then, optimal disturbance reject FOPID controller design is obtained as

$$C(s) = 49.9128 + 18.3780s^{0.7535} + \frac{32.1662}{s^{0.5358}} \quad (18)$$

Figure 15 shows a Bode diagram of the designed control system. The phase margin is indicated with a thick orange solid line in this figure. The phase margin value for the designed system is obtained as $\varphi_m = 2.1027$. Figure 16 shows step responses of the proposed FOPID controller and Tabatabaei et al.'s optimal FOPID controller[38]. Figure 17 shows a close view of Figure 16 in order to compare step performances. The figure indicates the improvement of the step response performance by means of the proposed FOPID controller in this example.

Figure 17 shows the RDR spectrum of the FOPID control systems. The proposed FOPID controller provides a superior RDR performance, and this indicates that the proposed FOPID controller presents a better disturbance rejection control performance compared to the Tabatabaei et al.'s FOPID controller. To improve setpoint control performance, authors used 2DOF setpoint filter FOPID control system (See Figure 1) in disturbance rejection control simulations for the proposed FOPID controller. The setpoint filter is configured as

$$F(s) = \frac{1}{3s+1} \quad \text{that avoids the high overshoot step response}$$

when settling to the setpoint 1. (Tabatabaei et al.'s FOPID control system does not include a set-point filter)

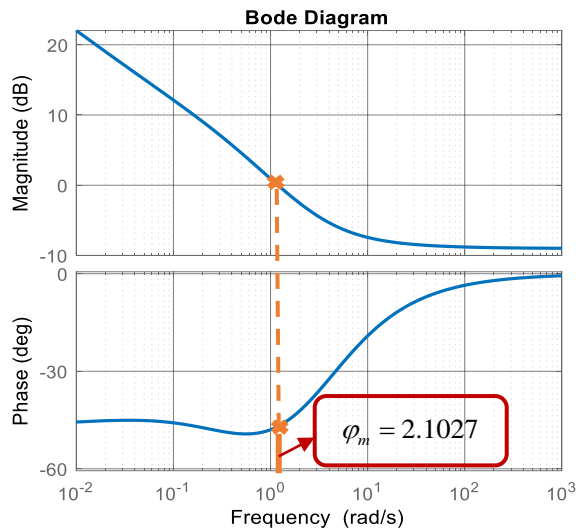


Fig.15. Bode diagram of the proposed control system

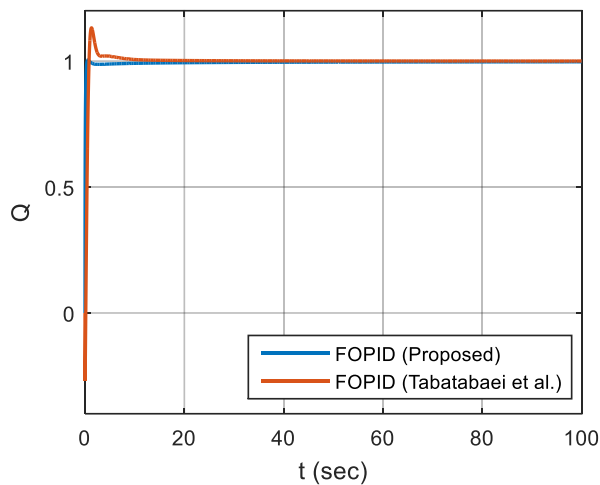


Fig.16. Step responses of the control systems

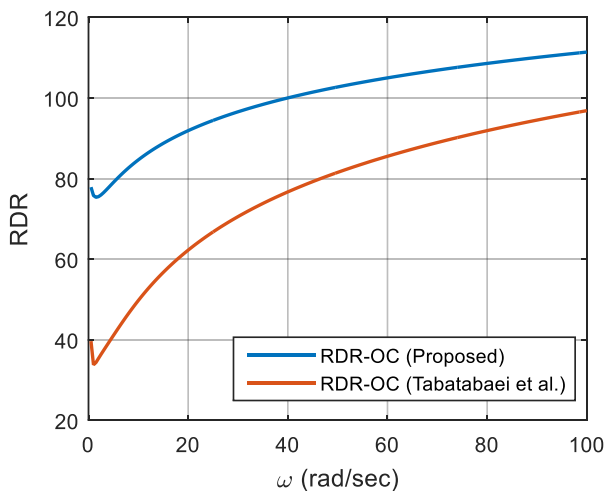


Fig.17. The RDR spectrums of the control systems

controllers. An additive step disturbance was applied to the control system at 40 sec simulation time. Figure 19 shows a close view of disturbance rejection control of both controllers.

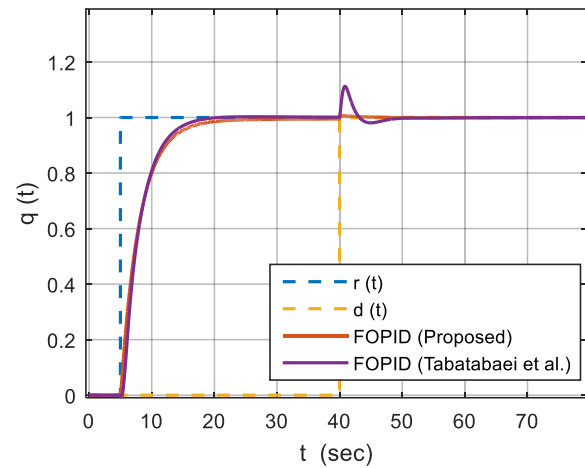


Fig.18. Step responses and disturbance responses of the FOPID controllers in case of an additive step disturbance insertion into the plant input at 40 sec

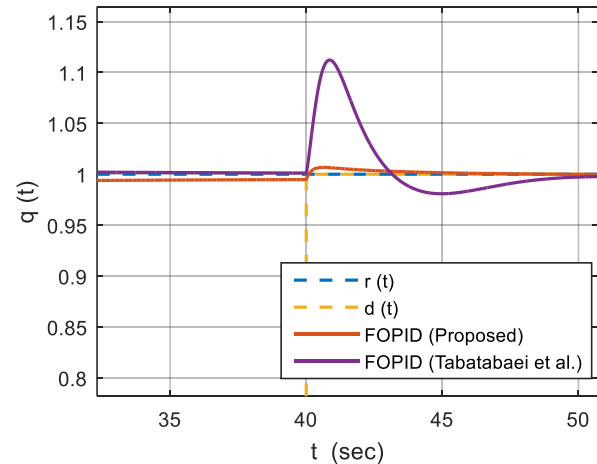


Fig.19. A close view of step responses in Figure 18

Figure 20 shows system responses when a sinusoidal input disturbance with the amplitude of 1 and a frequency of 3.14 rad/sec was inserted into the system. Figure shows that the disturbance rejection performance in this frequency is improved by the proposed FOPID control system.

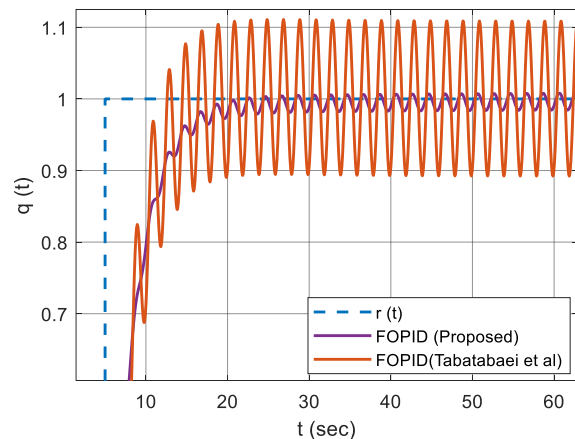


Fig.20. Sinusoidal disturbance responses of the control systems

Figure 18 shows step disturbance response of the

V. CONCLUSIONS

In summary, this study demonstrated a software implementation of the disturbance reject FOPID controller design methodology that was based on the RDR index maximization subject to a soft constraining of the phase margin. Thus, a difficult mathematical optimization problem can be effectively solved by using soft computing techniques.

The proposed control system performance was compared with other optimal FOPID system design methods. Simulation results reveal the improvement of the step response and the disturbance rejection performance compared to other optimal tuning methods. Future work can address the addition of more design constraints by using soft computing routines in order to solve more sophisticated design problems.

References

- [1] K.B. Oldham, J. Spanier, "Fractional Calculus." Academic Press Inc., 1974.
- [2] R. Hilfer, "Application of Fractional Calculus in Physics." World Scientific, 2000.
- [3] Y.Q. Chen, T. Bhaskaran, and D. Xue, "Practical tuning rule development for fractional order proportional and integral controllers." *J. Comput. Nonlinear Dyn.*, vol. 3, 2, 2008, pp. 214031–214038.
- [4] W. Jifeng, L. Yuankai, "Frequency domain Analysis and Applications for Fractional Order Control System." *IOP, Journal of Physics: Conf.* 13, 2005, pp. 268–273.
- [5] D. Xue, Y.Q. Chen, "A Comparative Introduction of Four Fractional Order Controllers." *Proceedings of the 4th World Congress on Intelligent Control and Automation*, Shanghai, P.R. China, 2002, pp. 3228–3235.
- [6] G.F. Franklin, J.D. Powell, A.E. and Baeini, "Feedback Control of Dynamic Systems." Addison-Wesley, Reading, MA, 1986.
- [7] K. Ogata, "Modern Control Engineering." 2nd edn. Prentice-Hall, Englewood Cliffs, NJ, 1990.
- [8] J.L. Chang, "Robust output feedback disturbance rejection control by simultaneously estimating state and disturbance." *J Control Sci Eng*, 2011, pp. 1–13.
- [9] Z. Vukic, O. Kuljaca, "Lect PID Controllers.", Vol. 23, 2002.
- [10] D. Vrancic, S. Strmčnik, J. Kocijan, P.B. de Moura Oliveirac, "Improving disturbance rejection of PID controllers by means of the magnitude optimum method." *ISA Trans*, Vol. 49, 1, 2010, pp. 47–56.
- [11] J.M.E. Vandeursen, J.A. Peperstraete, "Model-based and PID controllers for disturbance rejection in processes with time delay: a comparison." *ISA Trans*, Vol. 35, 3, 1996, pp. 225–236.
- [12] F. Fayaz and G.L. Pahuja, "Disturbance rejection based fractional order PID controller for system performance improvement of hybrid power system." *Optimal Control Applications and Methods*, 2022.
- [13] P. Selvaraj, O.M. Kwon, & R. Sakthivel, "Disturbance and uncertainty rejection performance for fractional-order complex dynamical networks." *Neural Networks*, Vol. 112, 2019, pp. 73–84.
- [14] F.N., Deniz, C., Keles, B.B., Alagoz, and N. Tan, "Design of fractional-order PI controllers for disturbance rejection using RDR measure." In *International Conference on Fractional Differentiation and Its Applications 2014 (ICFDA 2014)*, 2014.
- [15] B.B. Alagoz, F.N. Deniz, C. Keles, and N. Tan, "Disturbance rejection performance analyses of closed loop control systems by reference to disturbance ratio." *ISA Transactions*, Vol. 55, 2015, pp. 63–71. doi:10.1016/j.isatra.2014.09.013.
- [16] A. Biswas, S. Das, and A. Abraham, and S. Dasgupta, "Design of fractional-order PID controllers with an improved differential evolution." *Engineering Applications of Artificial Intelligence*, Vol. 22, 2009, pp. 343–350.
- [17] M. Zamani, M. Karimi-Ghartemani, N. Sadati, M. Parniani, "Design of a fractional order PID controller for an AVR using particle swarm optimization." *Control Engineering Practice*, Vol. 17, 2009, pp.1380–1387.
- [18] G. Zhe, X. Liao, "Rational approximation for fractional order system by particle swarm optimization," *Nonlinear Dynamics*, Vol. 67, 2, 2012, pp. 1387–1395.
- [19] M. Micev, M. Čalasan, & D. Oliva, "Fractional order PID controller design for an AVR system using Chaotic Yellow Saddle Goatfish Algorithm." *Mathematics*, Vol. 8, 7, 2020, 1182.
- [20] D. Izci, S. Ekinci, & B. Hekimoğlu, "Fractional-order PID controller design for buck converter system via hybrid Lévy flight distribution and simulated annealing algorithm." *Arabian Journal for Science and Engineering*, 2022, pp. 1–19.
- [21] I. Paducel, C.O. Safirescu, & E.H. Dulf, "Fractional Order Controller Design for Wind Turbines." *Applied Sciences*, Vol. 12, 17, 2022, 8400.
- [22] A. Richards, "Fast Model Predictive Control with Soft Constraints." *European Control Conference (ECC)* Julay, Zürich, Switzerland, pp. 17–19.
- [23] J.C. Regin, "Using hard constraints for representing soft constraints, *International Conference on AI and OR Techniques in Constraint Programming for Combinational Optimization Problems*." Springer, Berlin, Heidelberg, 2011, pp. 176–189.
- [24] P. Meseguer, F. Rossi, and T. Schiex, "Soft Constraints, *Foundation of Artificial Intelligence*." Elsevier, 2006, pp. 281–328 (Soft global constraining page 312).
- [25] M. Chandrasekaran, M. Muralidhar, C.M. Krishna, U.S. Dixit, "Application of soft computing techniques in machining performance prediction and optimization: a literature review." *The international journal of Advanced manufacturing Technology*, Vol. 46, 5, 2010, pp. 445–464.
- [26] S. Rahnamayan, H.R. Tizhoosh, M.M.A. Salama, "Opposition versus randomness in soft computing techniques," *Applied Soft Computing*, Vol. 8, 2008, pp. 906–918.
- [27] B.B. Alagoz, N. Tan, F.N. Deniz, C. Keles, "Implicit disturbance rejection performance analysis of closed loop control systems according to communication channel limitations." *IET Control Theory & Applications*, Vol. 9, 17, 2015, pp. 2522–2531.
- [28] A. Ates, B.B. Alagoz, C. Yeroglu, et al., "Disturbance rejection FOPID control of rotor by multi-objective bb-bc optimization algorithm." In: *ASME 2017 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, 13th ASME/IEEE International Conference on Mechatronic and Embedded Systems and Applications Cleveland, Ohio, USA, Vol. 9, 6–9 (August 2017).
- [29] S. Tufenkci, B. Senol, B.B. Alagoz, & R. Matuš, "Disturbance rejection FOPID controller design in v-domain." *Journal of Advanced Research*, Vol. 25, 2020, pp. 171–180.
- [30] N. Ozbey, C. Yeroglu, B.B. Alagoz, N. Herencsar, A. Kartci, & R. Sotner, "2DOF multi-objective optimal tuning of disturbance reject fractional order PID controllers according to improved consensus oriented random search method." *Journal of Advanced Research*, Vol. 25, 2020, pp. 159–170.
- [31] Q. Jin, Y. Shi, Q. Liu, M. Chu, & Y. Zhang, "Graphical robust PID tuning for disturbance rejection satisfying multiple objectives." *Chemical Engineering Communications*, Vol. 205, 12, 2018, pp. 1701–1711.
- [32] D. Goldberg, J. Holland, "Genetic Algorithms and Machine Learning." *Mach. Learn.*, Vol. 3, 2–3, 1988, pp. 95–99.
- [33] K.F. Man, K.S. Tang, ve S. Kwong, "Genetic Algorithms." Springer Publishing, 1999.
- [34] D.E. Goldberg, "Genetic Algorithms in Search, Optimization and Machine Learning." Addison-Wesley Publishing Company, 1989.
- [35] A.R. Tavakolpour, I.Z. Mat Darus, O. Tokhi, M. Mailah, "Genetic Algorithm-Based Identification of Transfer Function Parameters for a Rectangular Flexible Plate System." *Eng. Appl. Artif. Intel.*, Vol. 23, 8, 2010, pp. 1388–1397.
- [36] X. Dingyu, "Fractional-Order Control Systems: Fundamentals and Numerical Implementations," Walter de Gruyter, Berlin, Germany, Boston, MA, USA, 2017.
- [37] C. Zhao, D. Xue, & Y. Chen, "A fractional order PID tuning algorithm for a class of fractional order plants." In *IEEE International Conference Mechatronics and Automation*, Vol. 1, 2005, pp. 216–221.
- [38] M. Tabatabaei, R. Salehi, "Fractional order PID controller design based on Laguerre orthogonal functions." *International Journal of Dynamics and Control*, Vol. 5, 3, 2017, pp. 542–550.

APPENDIX

The matlab code of this design tool is given below. The objective function is costFunc(x):

```
function E = costFunc(x)
% x chromosome represents candidate controller
coefficients
kp=x(1);
ki=x(2);
kd=x(3);
lamda=x(4);
u=x(5);
% frequency setting for RDR index
walt=0;
wust=100;
dw=0.5;
w=walt:dw:wust;
% Calculation of logarithmic RDR index to
implement equations(5)and(2)
RDRdB=20*log10((kp+ki*cos((pi/2)*lamda)*w.^(-
lamda)+kd*cos((pi/2)*u)*w.^(u)).^2+(kd*sin((pi
/2)*u)*w.^(u)-ki*sin((pi/2)*lamda)*w.^(-
lamda)).^2);
% Calculation of minimum RDR index that
implements equation(9)
rdr=min(RDRdB);
%Plant parameters
a0=1;
b0=1;
b1=0.5;
alfa=0.9;
b2=0.8;
k=2.2;
%Phase margin specification
TargetPhaseMar=2*pi/3;
Pc=-pi;
% Coefficients of closed loop transfer
function
c0=a0*ki;
c1=a0*kp+b0;
c2=a0*kd;
c3=b1;
c4=b2;
beta0=0;
beta1=lamda;
beta2=lamda+u;
beta3=lamda+alfa;
beta4=k+lamda;
d2=a0*kd;
d1=kp*a0;
d0=ki*a0;
gamma2=(lamda+u);
gamma1=lamda;
gamma0=0;
% The software realization of the closed loop
transfer function
sistem=fotf([c4 c3 c2 c1 c0],[beta4 beta3
beta2 beta1 beta0],[d2 d1 d0],[gamma2 gamma1
gamma0]);
% The software realization of phase margin
calculations
X=bode(sistem,w);
Tw=squeeze(X.ResponseData);
Tw1=squeeze(Tw);
Mw=abs(Tw1);
Pw=angle(Tw1);
```

```
% Determination of crossover frequency by
using soft computing
% The following code detects zero crossing
for i=2:length(Mw)
if log10(Mw(i))*log10(Mw(i-1))<0
Wc=w(i);
Pc=Pw(i);
break;
end
end
%Phase margin according to equation (6)
Qp=pi+Pc;
% Implementation of optimization problem
according to soft constraint
% for phase margin specification
if abs(TargetPhaseMar-Qp)<0.01
E=(1/((rdr)^2)+0.01);
else
E=abs(TargetPhaseMar-Qp)*1e+10;
end
end
```

The code runs Matlab genetic algorithm for the objective function *costFunc(x)*:

```
clc;
clear all;
close all;
ObjectiveFunction = @costFunc;
nvars = 5;
LB = [1 1 1 0.5 0.3]; % Lower bound
UB = [50 50 50 2 2]; % Upper bound
opts =
gaoptimset('PlotFcn',{@gaplotbestf,@gaplotbest
indiv});
[x,fval,exitFlag,Output,population,scores] =
ga(ObjectiveFunction,nvars,[],[],[],[],LB,UB,[
],opts);
fprintf('****Optimization completed ****\n')
fprintf('kp : %f \n',x(1));
fprintf('ki : %f \n',x(2));
fprintf('kd : %f \n',x(3));
fprintf('lampda : %f \n',x(4));
fprintf('mu : %f \n',x(5));
fprintf('Minimum value of objective function :
%f \n',fval);
```

(To run this code for *fotf()* function, *fotf toolbox* should be placed in the same folder.)

BIOGRAPHIES



SEVILAY TUFENKCI graduated from Selçuk University Computer Engineering Department in 2017. She graduated her MSc program in Inonu University, department of Computer Engineering in 2019 and is continuing her PhD program in the Computer Engineering Department of Inonu University. Her research interests are

control systems, metaheuristic optimization and reinforcement learning methods.



BARIS BAYKANT ALAGOZ, received bachelor's degree in Electronics and Communication Engineering from Istanbul Technical University. He received the Ph.D. degree in Electrical-electronics Engineering from Inonu University in 2015. His research interests are computational intelligence, intelligent systems, control systems,

simulation and modeling.



CELALEDDİN YEROĞLU, received the BSc degree in electronics engineering from Hacettepe University, in 1991, and the Master and PhD degree from the Department of Computer Engineering, Trakya University, in 2000. He received second PhD degree in 2011 from the Department of Electrical and Electronics

Engineering, Inonu University. Since 2009, he has been working with the Computer Engineering Department, Inonu University. His research interests include intelligent systems, control theory and applications, simulation and modeling of networks.



BILAL SENOL received the B.S. and M.S. degrees in electrical and electronics engineering, and the Ph.D. degree in computer engineering from İnönü University, Malatya, Turkey, in 2009, 2011, and 2015, respectively. He is currently working as an Associate Professor with the Department of

Computer Engineering, İnönü University. His research interests include fractional-order analysis and controller design. He also works related to computer-aided design and user-friendly interfaces for system analysis. Lately, he is having researches on the classification of medical images.

Simple Design and Implementation of Two-Way Communication System through UAV

Yılmaz Kalkan*, Osman Avcı, Tolga Ulutaş, Engin Can Akar, Barış Köksal

Abstract—In this paper, an Unmanned Air Vehicle (UAV) which can carry and communicate with an Unmanned Ground Vehicle (UGV - Rover) is designed and realized. In this context, one UAV and one Rover are designed separately. The Rover is designed much smaller than the UAV as it is carried by the UAV and it can make its own movement after leaving the UAV near the area where it will be operated. At this stage, the carrier UAV acts as a communication relay, providing communication between the central control station and the Rover. Using nRF24L01 transceiver modules on each of the central control station, carrier UAV and the Rover, communication between them is achieved via programming as being master or slave. The relay communication system is designed on Serial Peripheral Interface (SPI) protocol of nRF24L01 transceiver modules. The Rover is also equipped with a camera on it to record videos and to take pictures. Both the UAV and the Rover were designed especially in order to reduce the energy consumption of them. Hence the Rover is designed as small in size and light as possible. By taking into account the equipment, batteries and motors that the vehicles will carry, the best design has been tried to be implemented in order to increase their ranges and reduce energy consumptions. It is possible to use it in the defence industry for military purposes, in places where bomb detection or long-range access is difficult with a single UAV, where land access is required.

Index Terms—Relay Communication, nRF24L01 Module, UAV, UGV, SPI.

I. INTRODUCTION

WITH THEIR high mobility and low cost, Unmanned Aerial Vehicles (UAVs), also commonly known as drones or remote-controlled aircraft, have found wide application in the last few decades [1]. Historically, drones were primarily used in the military, mostly deployed in enemy territory to reduce pilot casualties. With continued cost reduction and device miniaturization, small drones (typically not exceeding 25 kg) are now more easily accessible to the public. Therefore, numerous new applications have emerged in civil and commercial areas

• **Yılmaz Kalkan**(*corresponding author) is with the Department of Electrical and Electronics Engineering, Engineering Faculty, Aydın Adnan Menderes University, Aydın, 09010 TURKEY e-mail: yilmaz.kalkan@adu.edu.tr

• **Osman Avcı** is with the Department of Electrical and Electronics Engineering, Engineering Faculty, Aydın Adnan Menderes University, Aydın, 09010 TURKEY

• **Tolga Ulutaş** is with the Department of Electrical and Electronics Engineering, Engineering Faculty, Aydın Adnan Menderes University, Aydın, 09010 TURKEY

• **Engin Can Akar** is with the Department of Electrical and Electronics Engineering, Engineering Faculty, Aydın Adnan Menderes University, Aydın, 09010 TURKEY

• **Barış Köksal** is with the Department of Electrical and Electronics Engineering, Engineering Faculty, Aydın Adnan Menderes University, Aydın, 09010 TURKEY

Manuscript received May 27, 2022; accepted Dec 22, 2022.
DOI: 10.17694/bajece.1115408

with typical examples including weather monitoring, wild-fire detection, traffic control, cargo transportation, emergency search and rescue, communications relaying etc. [2], [3]. Many different vehicles (another drone or Unmanned Ground Vehicles (UGV)) can work simultaneously with these developing drones. They can act together in different missions as military or civilian and can perform tasks in cooperation. Monitoring and inspection tasks can be done with the drone, but in more sensitive missions the drone will not provide maximum benefit [4], [5]. More sensitive tasks can be performed with more sensitive vehicles like UGVs that can work with drones [6], [7], [8]. The communication with UGVs can be done both bilaterally and via drone. It means, the UGV can be guided autonomously by both human and drone [9], [10]. In some instances, multiple tools can work together to perform a task. No matter how much companies use this technology for military purposes, interaction between ground vehicles and drones will be seen more frequently in the following years and many different examples will be seen in the coming years.

The paper is organized as follows. In section 2 problem is defined clearly. Section 3 explains the design steps of both UAV and the Rover. Two-way communication between the units are given in section 4 and section 5 concludes the paper.

II. PROBLEM DEFINITION

The main purpose of this paper is to design an UGV (Rover) and an UAV that carries the Rover. Also their communications with each other and with the central control station problem is solved. This design can be used for the following purposes;

- 1) to be able to use it for applications that cannot be realized from the air and must be carried out from the ground.
- 2) to be able to collect data in a confidential way with the advantage of its small size in narrow and difficult regions to enter.
- 3) being able to deploy the Rover in longer distances with the help of communication established in areas where the UAV's range will not be enough.

The carrier UAV will pick up the Rover from its current position to move it to the desired location. When the carrier UAV arrives at its planned location, it lands and then releases the Rover. The carrier UAV begins to act as a communication relay so that it can communicate with the Rover and the central control station. After communication link is established, commands from the central control station pass over the carrier UAV and are transmitted to the Rover. The Rover performs the required missions and is taken back by the carrier UAV

and is returned back to the safe zone. This case is illustrated in Figure 1 and Figure 2.



Fig. 1. Realization of Project.

While the Rover performs the planned mission, the UAV operates in standby mode, consumes less energy and provides the communication between Rover and central control station. The UAV should be especially capable of carrying the Rover and be supported with equipment that can communicate with central control station during its flight. Moreover, Rover and the UAV must be designed to provide maximum flight time and consume less energy as possible. Details of the designs are more clearly explained in the design sections. The most important problems are to implement the two-way communication between the central control station-UAV-Rover and to design optimum UAV and Rover in terms of energy consumption and operation range.

Some mathematical calculations for weight lifting of the UAV and for motor and other device selections are included in the next sections.

III. DESIGN OF UAV

The 3 dimensional (3D) frame of the UAV was designed in SolidCad completely within the aerodynamic requirements. The UAV main body was determined to be an octocopter to fit for our purposes as shown in Figure 3 [11].

The main material of the UAV was determined as aluminium due to strong and easy-shaped characteristics. Aluminium



Fig. 3. UAV's 3D Frame Drawing In Computer.

materials were cut to the specified dimensions and the octocopter frame was welded to each other. The use of aluminium material normally provides a disadvantage in weight compared to carbon fibre material. However, this situation has been turned into a completely advantageous situation with the original design. The aluminium octocopter UAV, which we have created with our own unique design, provides 700 grams of lightness compared to normal carbon fibre frames. The combination of aluminium parts is provided by welding. Welding the arms of the frame that are separated to each other with the straight aluminium pipe in the middle, has also eliminated the strength problem. In this way, the UAV was created entirely with novel and original design. At the same time, the lightness provided has given us a positive advantage in lifting weights.

A. Thrust and Performance Calculations

The following calculations were made for motor and thrust calculations, respectively. As a result, compatible parts were determined for the purpose of the project. The total flight weight of the UAV (together with own weight and load) depends on the performances of the engines as below, program.

$$W_t = 2.19 \times T_{tp} \quad (1)$$

where W_t is the total weight (flight weight) of the UAV and T_{tp} is the total performance of the engines. After design is finished and the total weight of the UAV and the load is determined, by using Eq. 1 total engine performance required for this design can be determined and then each engine can be chosen easily. On the other hand, to calculate total weight of the drone, we need to determine engine due to the requirement of their weights. Then we need to check the selected engines satisfy the requirement given in eq.(1). The list of the selected parts of the UAV and their weights can be seen in Table I.

As it can be seen from Table 1, the total weight of the main components of the UAV was calculated as 3868 grams. This value is the total weight of the UAV when carrying the maximum load. By using the Eq.1 and the flight weight calculated in Tab. I, the thrust value of the motor can be calculate as [12];

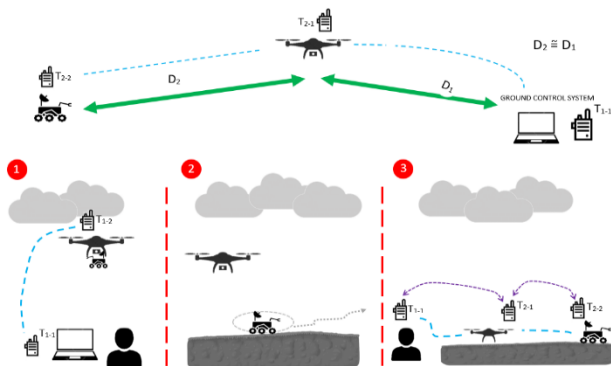


Fig. 2. Communication Schematic of Project.

TABLE I
WEIGHT OF UAV BODY EQUIPMENT.

Tool	Weight(gram)
Aluminum frame	680
Undercarriage body	224
Sunnysky X4108S brushless motor (8 pcs)	110 X 8 = 880
1245 propellers (8 pcs)	10 X 8 = 80
Battery (24 pcs)	48,5 X 24 = 116
Load (Rover)	840
Total	3868 g

$$3868 \times 2,19 = 8470,92 \text{ g} \quad (2)$$

According to the result above, 8470,92 g is the total thrust required by the motors. Due to octocopter case, 8 motors are used in the UAV design. Thrust force for each motor can be calculated as,

$$8470,92 \div 8 = 1058,87 \text{ g} \quad (3)$$

Each motor must apply 1058.87 grams of thrust. When the motors are compared in these standards, the selected engine (Sunnysky X4108S brushless motor) is found suitable for use in our design. The technical parameters of the selected motor is listed as in Tab. II.

TABLE II
TECHNICAL TABLE OF MOTOR [13].

Propelller (inch)	Voltage (V)	Current (A)	Thrust (g)	Power (W)	Efficiency (g/W)
		1	250	22.2	11
		2	430	44.4	9
		3	590	66.6	8
		4	720	88.8	8
		5	840	111	7
APC 1238	22.2	6	960	133.2	7
		7	1060	155.4	6
		8	1160	177.6	6
		9	1250	199.8	5
		10	1330	222	5
		11.7	1520	259.74	5

By looking at the Tab. II, we can see this motor can generate thrust from 250g up to 1520g. Our required maximum Thrust value for each motor is 1058,87g, therefore this motor can supply more power than required. It means these motors can be used safely to carry Rover. Our battery unit will supply 12A of current to the engines, on the other hand the nearest upper value of 1058,87 g is 1060g in table 2. Hence, it is required to check Thrust value of the motor for desired current value. Looking at the Table 2 at 1060g Thrust value of the motor, the current drawn by the motor should be checked. This value is shown in the table as 7 A. The thrust value required for the UAV was calculated as approximately 1058 grams. Hence selected motor is suitable for our design and each motor is drawn 7A current in full load. These values have been taken into account by the technical analysis results on the 12 inch

propeller. When the weight is added, the thrust value of the motor increases in the technical analysis results by increasing the propeller size. These results show that it can be solved by increasing the propeller size within the addition of parts to the UAV.

B. Battery Selection

Lithium ion battery is determined as the battery of the both UAV and the Rover. The reason for using a Lithium ion battery instead of a lithium polymer battery is that lithium ion has a high power density. Sony VTC 6 battery is used as battery. The output parameters of the battery were effective in this battery selection. Because the UAV needs a lot of firing power when taking off with its full load. This requirement has thus been solved. Each battery provides 3.7V and 3000mAh. Batteries were riveted to each other in the form of 6SP4 package. When the batteries are fully charged, a value of 22.2 Volts was obtained as output power. As a result, 12000mAh value and 22.2V values were provided by using 24 batteries in our design. On the other hand, as it is explained in the previous sub-section, 7A current will be enough. The advantage of the upper current value will be given in the next sub-section.

C. Calculation of Flight Time

The current value of the designed battery is 12000 mAh = 12 Ah. In order to calculate the flight time provided by the battery to the UAV, the following calculation steps are performed [14].

$$12A \times 60minutes = 720A \times minutes \quad (4)$$

Usually full engine efficiency cannot be achieved due to the losses. Universally 80% efficiency is achievable value for drone design. Hence,

$$720A \times minutes \times 0.8 = 576A \times minutes \quad (5)$$

The thrust value of the motor is based on 7 A. Because the value taken as a basis for weight corresponds to 7 A in this motor and we have total of 8 motors; current drawn can be calculated as;

$$7A \times 8 = 56A \text{ (total)} \quad (6)$$

and total flight time of the UAV with full load can be calculated by using Eq.5 and Eq.6 as;

$$(576A \times minutes) \div 56A = 10,29 \text{ minutes} \quad (7)$$

It means, the UAV provides approximately 10 minutes of flight with full load with our design and selected equipment. Certainly by standby mode, operation time of the UAV will be longer.

D. Flight Controller of UAV

An UAV includes many electronic part additional to mechanical parts. The most important electronic part of an UAV is the Flight control card. A Flight control card is the circuit board that manages the flight of the UAV. The flight controller's job is to control the power, or speed (Rotates per Minute - RPM), for each motor on the UAV in response to the information received from the UAV controller. It receives raw information from the UAV controller transmitted through the receiver and then makes the corresponding movement to the UAV's controller. Since UAVs have multiple motors, the flight controller's job is to take the information sent by the UAV controller and then act on it so that each individual motor receives the appropriate amount of power to execute the requested movement.

When a flight controller is set up well, the UAV pilot's commands sent via the UAV controller should correspond exactly to how the UAV moves in the air. Pixhawk flight control card was used because it fulfils the desired objectives and is also octocopter compatibility. It supports 8 electronic speed controllers (ESC) and 8 motors. Technical accuracy in terms of use and interface is higher quality. The avionics system wiring diagram of flight control card is shown in Fig.4. The camera, router and other additional systems are connected to another battery separately from the flight systems.

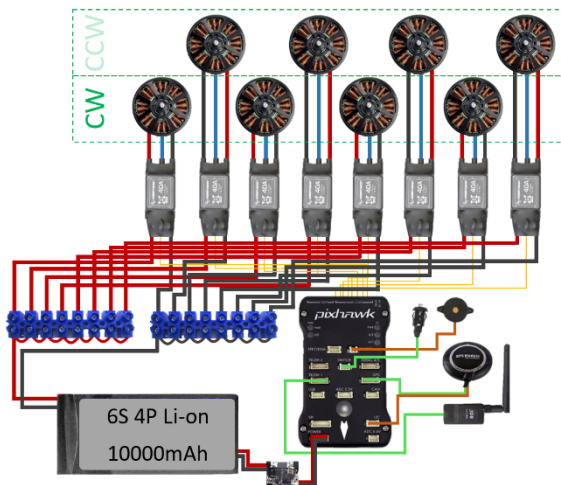


Fig. 4. The Avionics System Wiring Diagram.

E. PID Control Adjustments

To ensure a stable flight with and without load, the power of the each motor must be controlled instantaneously and the required adjustments must be done carefully and electronically as fast as possible. PID controllers are commonly used such adjustments reliably for UAV [15], [16], [17].

Adjusting to the UAV by using selected flight controller's (Pixhawk) user interface is a difficult process. The UAV needs to fly stable and hence, the most important factor is the tuning of the PID controller parameters. The Rover should not have any stability problems while flying with the UAV and weather conditions such as wind should be eliminated.

For adjustment, all controller parameters are set to be zero and maximum values should not be exceeded. The operating point where the control card's performance is the best should be determined. There are 3 axes being Pitch(P) - axis, roll(R) - axis and yaw(Y)-axis. On the P-axis, the UAV tilts forward or backward. On the R axis the UAV tilts right or left. On the Y axis, it rotates counterclockwise and rotates from above in the maintenance direction [19]. Usually P, D, I settings are made respectively. Then yaw control is tuned for stability.

The P controller is proportional to the angle of inclination of the motor power. As the motor output changes, the UAV is tried to be stabilized. If the P value is high, the UAV starts to swing and if the P value is too low, the UAV is flying unsteadily. It is the best to increase the P value until the UAV begins to oscillate than this value is set. D value controls the reaction time of UAV by processing the sensor data. If D value is low, it will not act on the rapid changes. If it is too high, it starts to oscillate when sending less signal [18]. The I controller is used to change the current drawn by motor hence the power of motor. This changes the deflection angle, provides overall correction and balances against P and D controllers on the UAV. I value was kept as small as possible considering the P control. The designed and produced UAV can be seen in Fig.5.



Fig. 5. The Designed UAV.

F. Design of Rover

Rover is very important for functional purpose, and it should be very light in all aspects as well. First, the frame suitable for the purpose is selected. The Rover was designed in a 3D drawing area to suit the intended purposes. A tracked all-terrain vehicle was chosen in terms of lightness and tough terrain conditions. A motor that is as light and functional as possible was chosen for the Rover. This motor is 6V 250 RPM dc motor. Each crawler arm was positioned to have 2 motors. A clipper is added to Rover design to give extra ability for example destroy the bomb for Rover. MG90s servo motor is used in this robot arm. In Fig.6 the Rover and clipper can be seen. All plastic materials (blue and white parts in Fig.6) were produced with 3D printer.

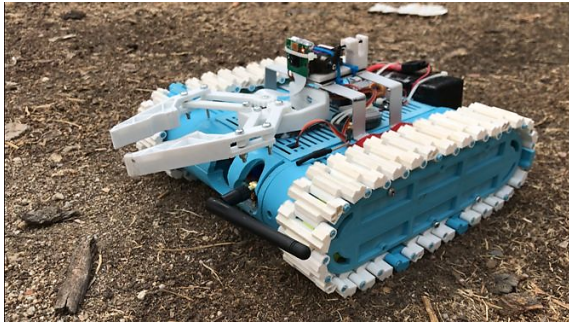


Fig. 6. The Designed UGV.

Rover includes 4 DC motors, L298N, Arduino UNO, nRF24L01 module, raspberry Pi, raspberry Pi camera module and 11.1 V battery. 4 motors are connected to L298n according to + and ground outputs. The nrf24L01 module was positioned on the Arduino UNO in line with the codes in the software. Appropriate connections to L298N have been made on Arduino Nano. Battery connected to L298N and raspberry PI module. Camera module is socketed to raspberry Pi. The Rover mechanism is completed as in Fig.6. Electrical connections can be seen in detail in Fig.7.

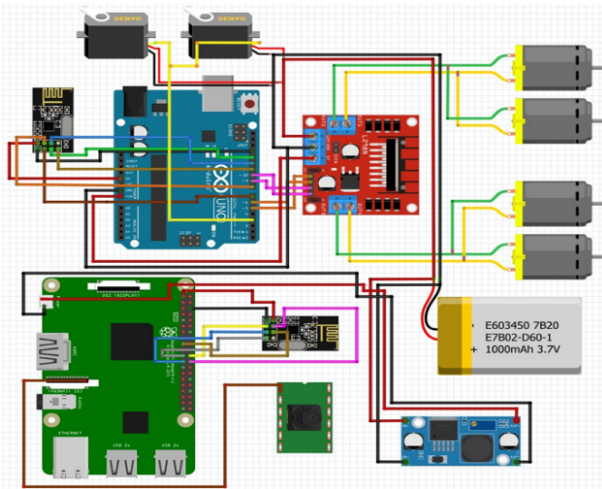


Fig. 7. Connections of Rover's Electronics Parts.

IV. COMMUNICATION SYSTEM

Nowadays, there are many protocols to be used for communication purposes. Concentrating on the necessary protocol according to the place and purpose of use always provides convenience. Therefore, the SPI Protocol was deemed the most appropriate for this project and studies were carried out on it.

A. SPI Protocols

SPI (Serial Peripheral Interface) is one of the synchronous serial communication types that Arduino supports. It is similar to I2C in terms of features and usage. It allows an Arduino to communicate with other Arduinos or sensors in a short distance. In SPI protocol, there is one Master device as in I2C. This device controls the connected peripheral devices. This is

a very necessary condition for the targeted project. There are four SPI lines, MISO (Master In Slave Out), MOSI (Master Out Slave In) and SCK (Serial Clock), CS (SS) (Chip Select) which connect to the master and peripheral devices.

As can be seen from the MISO and MOSI lines in Figure 8, the data lines are unidirectional in the SPI protocol, unlike I2C. In addition, peripheral devices (slave) do not need to have addresses [19]. Each peripheral device has a selection terminal. This terminal is called Slave Select (SS) terminal. The number of this line is equal to the number of peripheral devices activated [20]. For each device, a separate SS line emerges from the master device. Peripheral device with SS line LOW (0 volts) starts communication with master device [21]. SPI communication example is shown in Fig.8. As can be seen in this figure, there are SS outputs from the Master device as much as the number of peripheral devices. The master device pulls the SS pin of the peripheral device it wants to communicate with to the LOW (0 Volt) level [22].

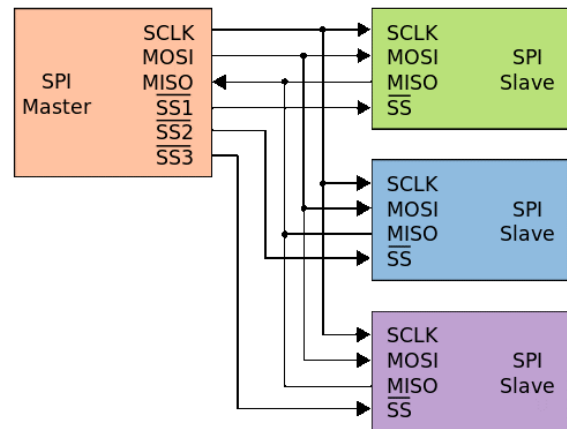


Fig. 8. An Example of SPI Communication Schema [22].

As can be seen in the Fig.8, the master part of the project is taken as the central control station. The antenna positioned on the UAV as the router and the Rover antenna, which is the end point, are operated as slave devices in our design.

B. nRF24L01 Communication Module

nRF24L01 is an inexpensive communication module that supports SPI communication protocol and has the necessary I/O pins, using the 2.4 GHz band. As mentioned before, SPI communication is an indispensable protocol for the Project. Therefore, this module is very advantageous for the project and it has been decided to use this module.

These modules are presented to the user in two types in the market. One of them does not contain an external antenna and is more suitable for close range use. The other antenna model, on the other hand, has few differences in design and has a high range in outdoor use. This module is called nRF24L01+ PA/LNA. The general structure of both models is shown in Fig.9.

Module with version without external antenna uses the built-in antenna and the other model has an SMA connector

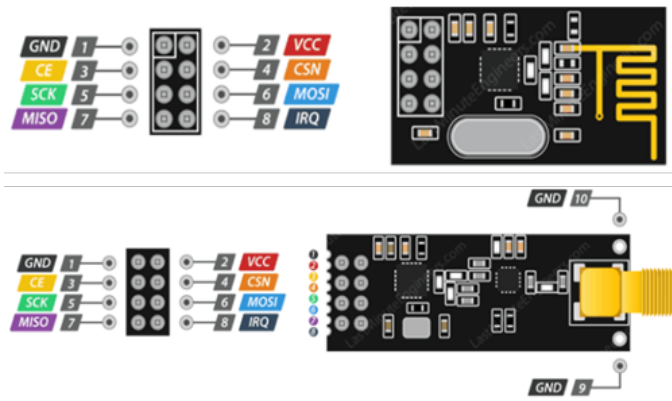


Fig. 9. General Structure of nRF24L01 Modules [23].

and a duck antenna. For this project, the use of the model with an external antenna from these two models was deemed appropriate. The nRF24L01+ transceiver module transmits and receives data on a specific frequency, called a Channel, and two or more transceiver modules must be on the same channel to communicate with each other [23]. This channel can be on any frequency of the 2.4 GHz ISM band, or more precisely between 2.400 and 2.525 GHz (2400 to 2525 MHz).

It occupies less than 1 MHz of bandwidth for each channel, providing 125 different channels with 1 MHz spacing. In this way, the module can use 125 different channels, giving the opportunity to have a network of 125 different and independent modules in one place. Channel 115 was used for communication within the scope of the project. This channel division is shown in Fig.10.

The nRF24L01+ provides an extra feature called Multi-ceiver. This means multi-transmitter is operated with single receiver. Each RF channel is logically divided into 6 parallel data channels, called Data Pipes. In other words, it is a logical channel in a physical RF Channel, just like a data pipe. Each data pipeline has its own configurable Data Pipe Address. This can be demonstrated as shown in Fig.11.

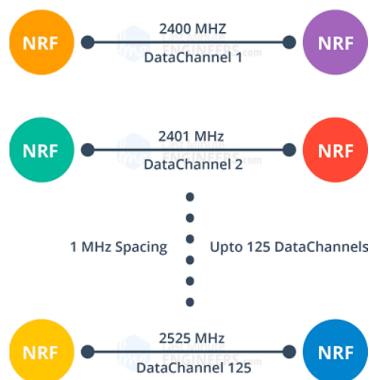


Fig. 10. Data Channel explanation of nRF24L01.

To explain the diagram above, the primary receiver is considered to act as a hub receiver, collecting information from 6 different transmitter nodes simultaneously. In this way, the Hub receiver can stop listening at any time and act as a



Fig. 11. Multiceiver Operation of nRF24L01.

transmitter. But during this process, it can only spawn one pipe/node at a time.

The nRF24L01+ transceiver module is designed to operate in the worldwide 2.4 GHz ISM frequency band, and the data transfer rate can be 250kbps, 1Mbps and 2Mbps. The operating voltage of the module is between 1.9V and 3.6V, but the good news is that the logic pins are 5 volt tolerant and can be easily connected to 5V without using a converter [24]. The module supports programmable output power. It consumes an incredible 12 mA at 0 dBm. And best of all, it consumes 26 μ A in standby mode and 900 nA in power off mode which are very low. Therefore, they are wireless devices for low-power applications. Also, it supports the advanced ShockBurst™ acceleration protocol [23].

C. End-Point (Rover) and Master (Control Station) Communication

The communication between the Rover Receiver, called End Point, and the Central Control Station is provided by the module on the UAV. The module on the UAV acts as a routing here. Therefore, it is named as Router. In short, the data from the Central Control Station was transferred to the endpoint via the router. Communication management is shown schematically in Fig.12.



Fig. 12. Explanation of Whole Communication System.

D. Control Station (Master) – UAV (Router) Communication

This part is one of the important communication parts of the project. This communication takes place by communicating with the Central Control Station and the nRF24L01 module

on the UAV, which helps to send the data and commands we give from the joystick to the UAV. This stage is the starting point of other communication, that is, the UAV acts as a bridge and transmits the parameter information it receives from the control station to the Rover. Conversely, the Rover transmits the collected data to the Central Control Station via the UAV. The electrical wiring diagram of the controller and controller of Rover is shown in Fig.13.

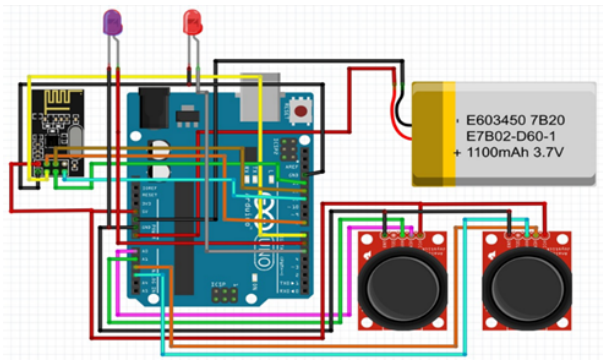


Fig. 13. Electrical Wiring of Controller.

The flowchart of the transmitter antenna module that acts as a transmission in the Central Control Station is shown in Fig.14.

E. UAV (Router) – Rover (End-Point) Communication

This connection between the UAV and the Rover, which is another important and main purpose of the project, in such cases as narrow areas where the UAV cannot reach the planned address, the range of the UAV alone is not enough and the battery is a problem, the desired operation can be carried out thanks to this communication. This communication is also provided with nrf24L01 modules. Thanks to this connection, the image and sound recording data collected by the Rover is first transmitted to the UAV and then transmitted to the Central Control Station via the UAV. Relevant codes are given in the conclusion section. The algorithm flowchart of the router antenna module that acts as a bridge network on the UAV and receive antenna module on the Rover is shown in Fig.15.

F. Communication Range Test

In the light of the designed algorithms, range tests were carried out in different scenarios. Two tests were conducted in areas with high building density (city) and in a completely open area with no buildings or obstructions. Usually the building or any wave absorbing structure is called communication interferences. Therefore, there was quite a lot of communication interference in the first test performed. For these reasons, substantially different results were obtained in the two cases. As it can be seen in Fig.17, the communication range from Central Control Station to UAV (as router) is 315m and the range between UAV and the Rover is 330 m. It is shorter than the range of open area as expected since there are a lot of interferences and obstacles between the communication modules.

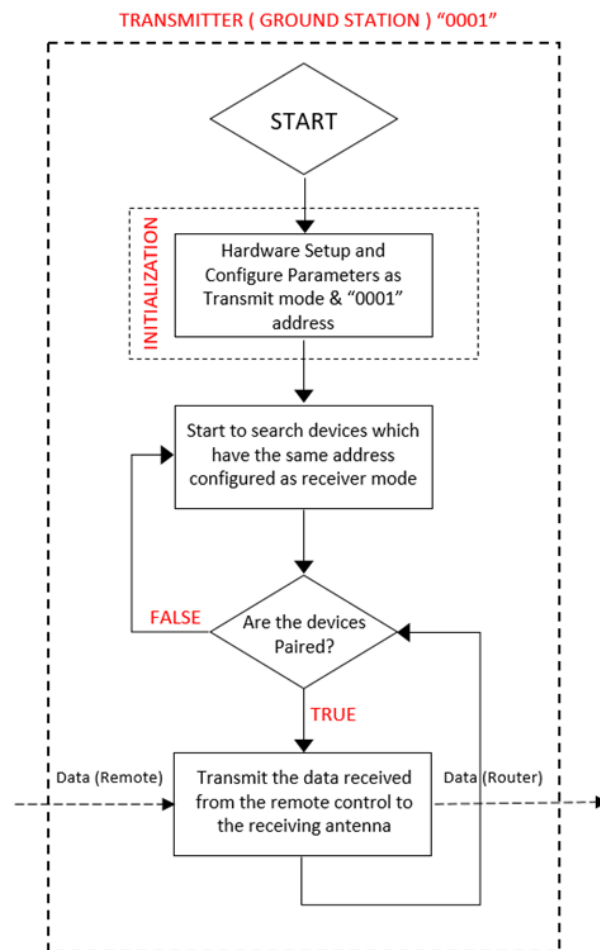


Fig. 14. Flowchart of Transmitter Algorithm.

The test results obtained in the open area have shown that the project can be used in practice and high ranges can be achieved with such an antenna in terms of budget. The open area test results are shown in Fig.18, and now it is clear that the communication ranges are increases as expected.

Now, the communication range from Central Control Station to UAV (as router) is 1480 m (instead of 315m) and the range between UAV and the Rover is 1520 m (instead of 330 m). We achieved the 5 times longer ranges in both. These results are summarized in Table III.

TABLE III
MAXIMUM OBTAINED COMMUNICATION RANGES BETWEEN THE UNITS.

	Ranges, m		
Region	CS - UAV	UAV - Rover	Total
Urban	315	330	645
Rural	1480	1520	3000

It has been seen that the designed system doubles the range between Central Control Station and the Rover which is the main idea of this paper. On the other hand, these are two-way communications between the units as well. The footage of camera on Rover is transferred to the Central Control Station

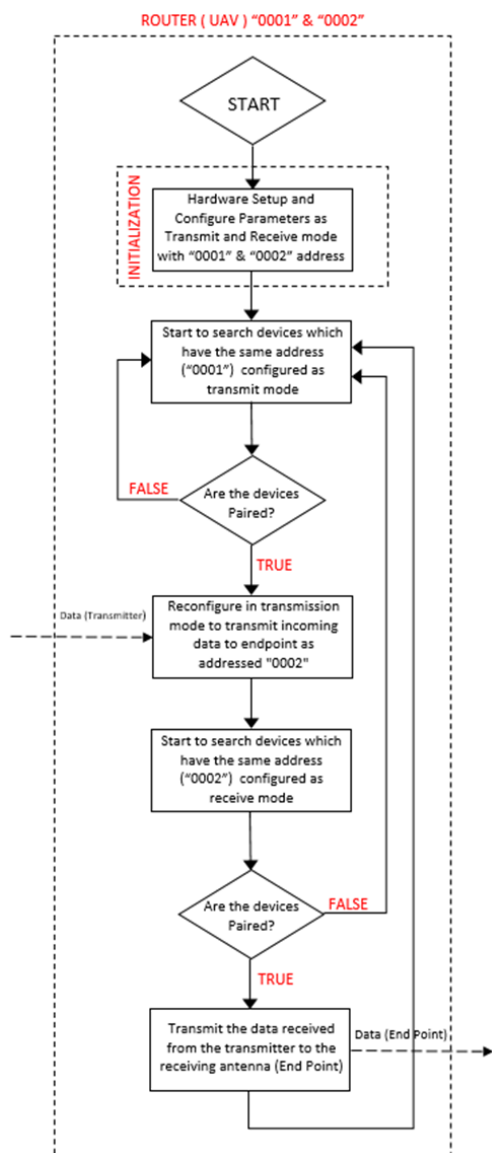


Fig. 15. Flowchart of Router Algorithm.

with some delay. Delay amount is proportional with the range and it is acceptable for critical missions.

V. CONCLUSION

In this paper, two-way communication system in between ground station and a ground vehicle over a UAV is designed and implemented with commercial and cheap microcontrollers, transceiver modules, flight controller, motors, sensors and 3D printer. The most difficult problem is the 2-way communication of the UAV as a router between the two modules (Central Control Station and Rover) at the same time without losing information. There are some delays in the 2-way communication between the units, but these delays can be further decreased with the selection of the superior microcontrollers and transceiver modules. Certainly this will increase the cost of the system.

Another important and critical problem to be solved is ensuring the stable flight for UAV with full load. Flight

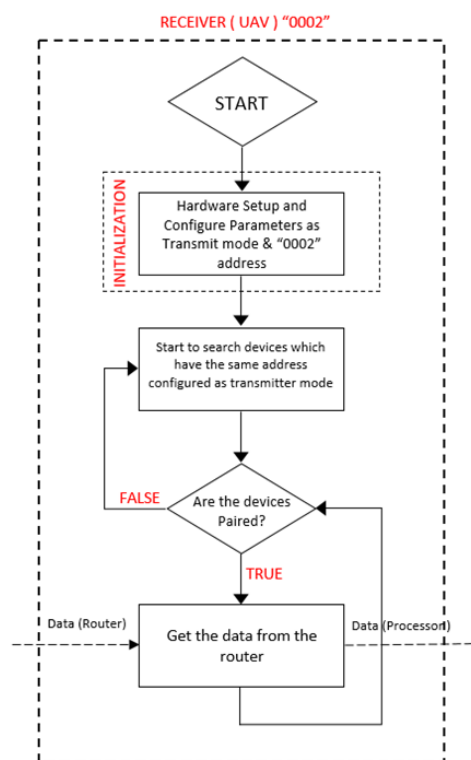


Fig. 16. Flowchart of Receiver Algorithm



Fig. 17. Urban Communication Test Result.

stabilization has been fully ensured and PID control algorithms have been developed successfully. The Rover, which will serve with the UAV, has been successfully implemented and fulfilled the tasks given in the real environment. With the established relay communication, many problems are solved. The release mechanism of Rover is fully provided. The UAV will leave the Rover exactly where the Rover will work. As stated before, the communication range has been tested under the influence of interference and without being under the influence of interference, and it is shown that communication range is doubled with this system.

The generated codes, circuit schematics and all required data to reproduce this project can be obtained from the following address;

<https://github.com/irisuavgithub/Spy-Drone-Project—Relay-Communication-System-with-nRF24L01->



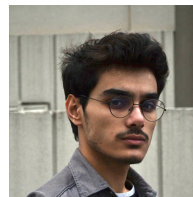
Fig. 18. Open Area (Rural) Communication Test Result.

REFERENCES

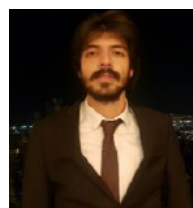
- [1] Y. Zeng, R. Zhang, and T. Lim, "Wireless communications with unmanned aerial vehicles: Opportunities and challenges," *IEEE Comm. Magazine*, vol. 54(5), pp. 36–42, 2016. [Online]. Available: <https://doi.org/10.1109/MCOM.2016.7470933>
- [2] B. Custers, *The Future of Drone Use: Opportunities and Threats from Ethical and Legal Perspectives*. T.M.C. Asser Press, 2016.
- [3] J. Lyu, Y. Zeng, R. Zhang, and T. Lim, "Placement optimization of uav-mounted mobile base stations," *IEEE Comm. Letters*, vol. 21(3), pp. 604–607, 2017. [Online]. Available: <https://doi.org/10.1109/LCOMM.2016.2633248>
- [4] M. Reinecke and T. Prinsloo, "The influence of drone monitoring on crop health and harvest size," in *In Proceedings of 1st International Conference on Next Generation Computing Applications*, 2017. [Online]. Available: <https://doi.org/10.1109/NEXTCOMP.2017.8016168>
- [5] Z. Duan, Y. Li, J. Wang, G. Zhao, and S. Svanberg, "Aquatic environment monitoring using a drone-based fluorosensor," *Appl. Phys. B*, vol. 125(108), pp. 36–42, 2019. [Online]. Available: <https://doi.org/10.1007/s00340-019-7215-y>
- [6] G. Quaglia, P. Cavallone, and C. Visconte, "Agri-q: Agriculture ugv for monitoring and drone landing," in *In Proceedings of the 4th IFToMM Symposium on Mechanism Design for Robotics*, 2019. [Online]. Available: <https://doi.org/10.1007/978-3-030-00365-4>
- [7] E. Harik, F. Guerin, F. Guinand, J. Brethe, and H. Pelvillain, "Uav-ugv cooperation for objects transportation in an industrial area," in *In Proceedings of the IEEE International Conference on Industrial Technology (ICIT)*, 2015.
- [8] A. Lakas, B. Belkhouche, O. Benkraouda, A. Shuaib, and H. Alasmawi, "A framework for a cooperative uav-ugv system for path discovery and planning," in *In Proceedings of the International Conference on Innovations in Information Technology (IIT)*, 2018. [Online]. Available: <https://doi.org/10.1109/INNOVATIONS.2018.8606028>
- [9] M. Saska, T. Krajník, and L. Pfeucl, "Cooperative μ uav-ugv autonomous indoor surveillance," in *In Proceedings of the International Multi-Conference on Systems, Signals & Devices*, 2012. [Online]. Available: <https://doi.org/10.1109/SSD.2012.6198051>
- [10] D. Galar, U. Kumar, and D. Seneviratne, *Robots, Drones, UAVs and UGVs for Operation and Maintenance*. CRC Press, 2020.
- [11] S. Kontogiannis and J. Ekaterinaris, "Design, performance evaluation and optimization of a uav," *Aerosp. Sci. Technol.*, vol. 29(1), pp. 339–350, 2013. [Online]. Available: <https://doi.org/10.1016/j.ast.2013.04.005>
- [12] Z. Benic, P. Piljek, and D. Kotarski, "Mathematical modelling of unmanned aerial vehicles with four rotors," *Interdiscip. Descr. Complex Syst.*, vol. 14, pp. 88–100, 2016. [Online]. Available: <https://doi.org/10.7906/indecs.14.1.9>
- [13] S. n.d., "Sunnysky x4108 brushless motor (datasheet)," <https://tr.aliexpress.com/item/32224922674.html>, Accessed December 2, 2021.
- [14] D. Myers, "Calculating flight time for unmanned aerial vehicles in the presence of obstacles and the incorporation of flight dynamics," in *MSc diss., State University of New York. ProQuest (AAT1482242)*, 2010.
- [15] A. Elbanna, T. Soliman, A. Ouda, and E. Hamed, "Improved design and implementation of automatic flight control system (afcs) for a fixed wing small uav," *Radioengineering*, vol. 27(3), pp. 339–350, 2018. [Online]. Available: <https://doi.org/10.13164/re.2018-3>
- [16] Q. Yuan, Y. Hu, C. Wang, and X. Ma, "3d beamforming for improving the security of uav-enabled mobile relaying system," *Radioengineering*, vol. 28(1), pp. 304–311, 2019. [Online]. Available: <https://doi.org/10.13164/re.2019-4>
- [17] B. Kamel, B. Yasmina, B. Laredj, I. Benamoumeur, and A. Zoubir, "Dynamic modeling, simulation and pid controller of unmanned aerial vehicle uav," in *In Proceedings of the 7th International Conference on Innovative Computing Technology, (INTECH)*, 2017.
- [18] V. Gomez, N. Gomez, J. Rodas, E. Paiva, M. Saad, and R. Gregor, "Pareto optimal pid tuning for px4-based unmanned aerial vehicles by using a multi-objective particle swarm optimization algorithm," *Aerospace*, vol. 7(6), pp. 604–607, 2020. [Online]. Available: <https://doi.org/10.3390/AEROSPACE7060071>
- [19] H. Keskin, "What is spi and how it works?" <https://herenkeskin.com/spi-nedir-ve-nasil-calisir/>, Accessed August 27, 2021.
- [20] M. Grusin, "Serial peripheral interface (spi)," <https://learn.sparkfun.com/tutorials/serial-peripheral-interface-spi/all>, Accessed August 28, 2021.
- [21] F. Baskoro, M. Rohman, and A. P. Nurdiansyah, "Serial peripheral interface (spi) communication application as output pin expansion in arduino uno," *Indonesian Journal of Electrical and Electronics Engineering (INAJEEE)*, vol. 3, pp. 34–40, 2020. [Online]. Available: <https://doi.org/10.26740/inajeee.v3n2.p63-69>
- [22] T. n.d., "Spi protokolü, geleceği yazanlar (turkish)," <https://gelecegiyazanlar.turkcell.com.tr/konu/egitim/arduino-401/spi-protokolu>, Accessed August 25, 2021.
- [23] L. M. E. n.d., "How nrf24l01 wireless module works & interface with arduino," <https://lastminuteengineers.com/nrf24l01-arduino-wireless-communication/>, Accessed August 22, 2021.



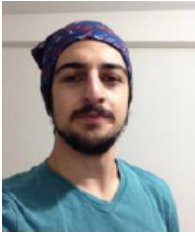
Yılmaz KALKAN was born in İzmir, Turkey, in 1979. He received the B.Sc. degree in Electronics and Communication Engineering in 2002 from Kocaeli University, Kocaeli, Turkey and Ph.D. degree in Electrical and Electronics Engineering from Middle East Technical University, Ankara, Turkey, in 2012. Currently, he is an associate professor at Electrical and Electronics Engineering department of Adnan Menderes University of Aydın, Turkey. His research interests are in signal processing, signal modelling and radar signal processing.



Osman Avcı was born in Rize, Turkey, 1998. He graduated from Electrical and Electronics Engineering department of Adnan Menderes University of Aydın, Turkey. His interests are in microcontrollers, power electronics and aviation electronics. He is currently working on UAVs in a company as an R&D Engineer.



Tolga Ulutaş was born in İzmir, Turkey, 1998. He is an undergraduate student at Electrical and Electronics Engineering department of Adnan Menderes University of Aydın, Turkey. His interests are embedded systems, control systems, analog electronics and multicopters.



Engin Can Akar was born in Iğdır, Turkey, 1998. He graduated from Electrical and Electronics Engineering department of Adnan Menderes University of Aydın, Turkey. His interests are IOT, power electronics and ROS.



Barış Köksal was born in Denizli, Turkey, 1998. He is an undergraduate student at Electrical and Electronics Engineering department of Adnan Menderes University of Aydın, Turkey. His interests are communications systems and software.

Single Image Dehazing based on Additive Wavelet Transform

Nur Huseyin Kaplan

Abstract—In this work, a single image dehazing method which uses the multi-scale products of additive wavelet transform as a prior is presented. In this method, first, the additive wavelet transform is applied to the hazy image to obtain its approximation and wavelet layers. Then the multi-scale products of the approximation and detail layers of the input hazy image are calculated. The multi-scale products of approximation and wavelet layers are summed up to obtain the proposed prior. Observations demonstrate that the proposed prior calculation keeps the detail information of the image, while detecting the haze. Using the proposed prior and commonly used hazy image model together, an efficient dehazing method is constructed. The comparisons between the proposed method and commonly used dehazing methods show that the proposed method has better dehazing performance than the traditional methods.

Index Terms—Dehazing, Defogging, Additive wavelet transform.

I. INTRODUCTION

IMAGING SYSTEMS tend to construct low contrast images especially in bad weather conditions like fog and haze. The low visibility of these image is dependent on the scattering and absorbing of the light by the particles within the atmosphere. Since, most of the automated systems and advanced image processing algorithms need higher contrast images, dehazing is an important problem for these applications.

Since, the amount of the haze depends on the distance between scene and the imaging device, the depth information of these images are not present. Therefore, image dehazing is a challenging problem because of the unknown depth information. The traditional enhancement methods, like histogram based methods are not capable of restoring the depth of the images [1], [2], [3].

In order to increase the performance of dehazing algorithms, multiple images of the same scene taken under different weather conditions are used [4], [5], [6]. These algorithms are effective for haze removal, however the need of the multiple scenes for input is generally not possible.

In recent years, many single image dehazing algorithms have been developed [7], [8], [9], [10], [11], [12], [13], [14], [15], [16]. These methods generally use a prior or assumption alongwith the common hazy image model.

Tan et al. [7] proposed a method to improve the local contrast of the hazy images under the assumption that haze free images have higher contrast. However, the haze free images are not always higher in contrast, this method distorts the color

information of the input images. Fattal [9] proposed a dehazing algorithm based on Independent Component Analysis (ICA). This method produces better results than Tan's method, however, the method is developed for only color images and does not work well for heavy hazed images.

He et al. [8] proposed to use the Dark Channel Prior (DCP) for dehazing. The prior calculates the darkest pixel within a patch, based on the observation that at least one pixel has low intensity for at least one color channel. Although being an effective method, it distorts the color, especially in the sky region and the time consumption is too high due to the soft matting process to reduce the effects caused by patches. In [14], the shortcomings of DCP have been investigated, and a dehazing method based on bright object handling is proposed.

Zhu et al. [13] has proposed a colour attenuation prior for image dehazing with a supervised learning process.

In recent years, dehazing methods based on decomposition [17] and fusion [18] strategies have also been proposed. Moreover, many learning algorithms [19], [20] are also used for dehazing purposes. Cai et al. [15] has proposed a trainable system (DehazeNet) to estimate the transmission with the help of convolutional neural networks.

Recently a multi-scale products prior based method (MP) has been proposed [16]. In this method, the input hazy image is decomposed by an undecimated Laplacian decomposition to obtain its detail and approximation subbands. Then, the multi-scale product of the approximation subband is used to calculate the prior. This prior is capable of removing the haze, as well as it does not distorts the color in the sky region. However, the method can not remove the haze especially in the mid regions. In addition, the method can only be applicable for color images.

Motivated by the results obtained in [16], an additive wavelet transform based method is proposed in this paper. The proposed method decomposes the input image to obtain the approximation and wavelet layers with additive wavelet transform. The proposed method takes into consideration both the approximation and wavelet layers to obtain the prior. Based on the proposed prior and the commonly used hazy image model, an effective dehazing algorithm is developed.

A large dataset of hazy image has been used to examine the proposed method. The resulting images demonstrate that the proposed dehazing algorithm produces images that have higher contrast and visibility. Moreover, the proposed method does not have the distortions of the traditional methods, as well as it is more efficient than the MSP method for mid regions.

Rest of the paper is as follows. An explanation of the additive wavelet transform is given in Section II. The proposed

Nur Huseyin Kaplan is with the Department of Electrical and Electronic Engineering, Engineering and Architecture Faculty, Erzurum Technical University, Erzurum, 25050 TURKIYE e-mail: huseyin.kaplan@erzurum.edu.tr
orcid ID: 0000-0002-4740-3259

Manuscript received Jun 08, 2022; accepted Dec 16, 2022. DOI: 10.17694/bajece.1127633

dehazing method is given in Section III. The comparisons of the proposed method with traditional methods are included in Section IV. The paper is concluded in Section V.

II. ADDITIVE WAVELET TRANSFORM

Additive Wavelet Transform (AWT) has a wide area of use in image processing applications such as pansharpening [21], because of its shift-invariance property. In this transform, first an approximation layer is obtained by filtering the input image (I) with a bicubic spline filter

$$I_1 = I * h_1 \quad (1)$$

where, I_1 is the first approximation layer and the bicubic spline filter h_1 can be given as:

$$h_1 = \begin{bmatrix} 1 & 4 & 6 & 4 & 1 \\ 4 & 16 & 24 & 16 & 4 \\ 6 & 24 & 36 & 24 & 6 \\ 4 & 16 & 24 & 16 & 4 \\ 1 & 4 & 6 & 4 & 1 \end{bmatrix} \quad (2)$$

The difference between the approximation layer and the input image gives the first wavelet layer.

$$W_1 = I - I_1 \quad (3)$$

For further decomposition, the first approximation layer is filtered again as

$$I_2 = I_1 * h_2 \quad (4)$$

Here, I_2 is the 2 nd level approximation layer, h_2 is the 2 nd level bicubic spline filter, which is obtained by doubling the size of h_1 by filling the gaps with zeros. The 2 nd level detail layer can be obtained as:

$$W_2 = I_1 - I_2 \quad (5)$$

To obtain the approximation and details layer of l th level, the equations are generalized as

$$I_l = I_{l-1} * h_l \quad (6)$$

$$W_l = I_{l-1} - I_l \quad (7)$$

The original image can be reconstructed by:

$$\hat{I} = I_L + \sum_{i=1}^L W_i \quad (8)$$

where L is the level of decomposition.

III. THE PROPOSED DEHAZING METHOD

Commonly used hazy image model can be given as [9], [16], [22], [23]:

$$I_m(x) = J_m(x)t(x) + A_m[1 - t_m(x)], m = R, G, B \quad (9)$$

where $I_m(x)$ and $J_m(x)$ are the m th band of the input hazy and haze free images, respectively. A_m is the m th band airlight, while $t_m(x)$ is the transmission of the m th band, which describes the unscattered data.

By using (7), $J_m(x)$ can be determined by using the following equation.

$$J_m(x) = \frac{I_m(x) - A_m[1 - t(x)]}{t(x)} \quad (10)$$

In order to obtain the haze free $J_m(x)$ image, the unknown airlight A_m and the transmission $t(x)$ should be estimated.

A. The Proposed Prior

In MP method described in [16], the prior has been obtained by the multi-scale products of the approximation subbands of the undecimated Laplacian decomposition. Moreover, a single prior is obtained for the color image.

In this work, the proposed prior is obtained by the use of both approximation and wavelet layers of the AWT. In addition, priors for every band of image is calculated separately.

In order to obtain the proposed prior, first the approximation and wavelet layers of the input hazy image are calculated by applying the AWT to the each color band of the image. The approximation layers are calculated as:

$$I_l^m = I_{l-1}^m * G_l^m \quad (11)$$

where $m = R, G, B$ is the color band of the image and $l = 1, 2, \dots, L$ is the level of decomposition. I_l^m is the l th level approximation layer of the m th band.

The wavelet layers are obtained by the following equation.

$$W_l^m = I_l^m - I_{l-1}^m \quad (12)$$

where W_l^m is the l th level wavelet layer of m th band.

Then, approximation multi-scale product for each band is calculated by multiplying adjacent approximation layers.

$$MSP_I^m = I_1^m I_2^m \dots I_L^m \quad (13)$$

and the wavelet multi-scale product for each image band is obtained by multiplying the adjacent wavelet layers.

$$MSP_W^m = W_1^m W_2^m \dots W_L^m \quad (14)$$

If a pixel value at a certain location is higher, corresponding pixel value at adjacent layer is higher too, and vice versa [24]. Multiplying adjacent layers to calculate multi-scale products may dilute redundant information and emphasize the important information of the image [24].

Finally, the proposed additive wavelet prior (AWP) for the m th band is calculated as:

$$AWP_m = \sqrt[L]{MSP_I^m + MSP_W^m} \quad (15)$$

In (15), the calculated wavelet and approximation multi-scale products are summed up. By this way, the priors calculated separately for each band, which makes the proposed method to be applicable for gray scale images as well. Including the wavelet layers in the summation makes the prior to keep detail information.

The flowchart of the calculation of proposed prior is given in Fig.1.

In order to determine the decomposition level, the effect of the levels on the prior is given in Fig.2.

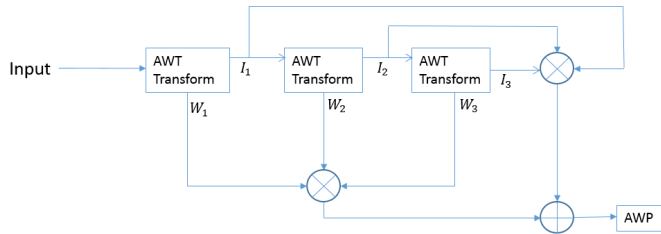


Fig. 1. Flowchart for the proposed AWP.

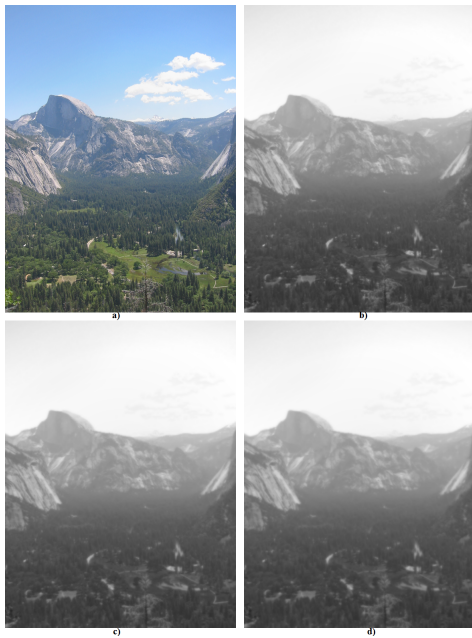


Fig. 2. a) Input hazy image, the proposed AWP obtained for b) 2 levels, c) 3 levels, and d) 4 levels of decomposition.

As seen in Fig.2, choosing the decomposition level higher than 2, the proposed prior tends to be blurry, which means that the prior is not able to keep important information of the image. Observing various hazy images, 2 levels of decomposition have been chosen.

Since, the proposed prior uses the approximation and wavelet layers together, the prior expected to keep both rough information and detail information of the input image. The proposed prior is used to estimate transmission, therefore by the use of this prior, a better transmission is expected to be obtained.

B. The Airlight and Transmission Estimates

The airlight estimation is equal to the mean of the brightest 0.1% pixels of the image as described in [8], [16].

For the estimation of the transmission, a similar way described in [8], [16] is followed. First, the hazy image is divided by the calculated airlight and proposed AWP of this normalized image is obtained. The calculated AWP is smoothed by the bicubic spline filter. The equation for transmission calculation can be formulated as:

$$t_m(x) = 1 - \omega AWP \left(\frac{I_m}{A_m} \right) * h_1 \quad (16)$$

where, ω is the amount of the haze to be removed. A transmission estimate with the proposed prior is given in Fig.3.

The dehazed images with different ω values are given in Fig.4.

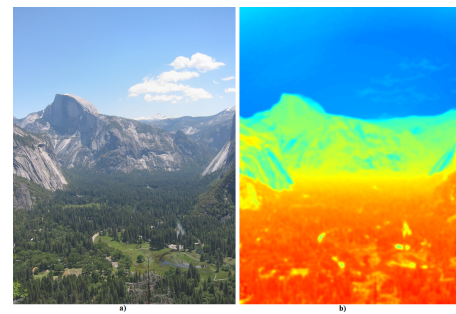


Fig. 3. a) Hazy image b) Transmission estimate with proposed AWP.

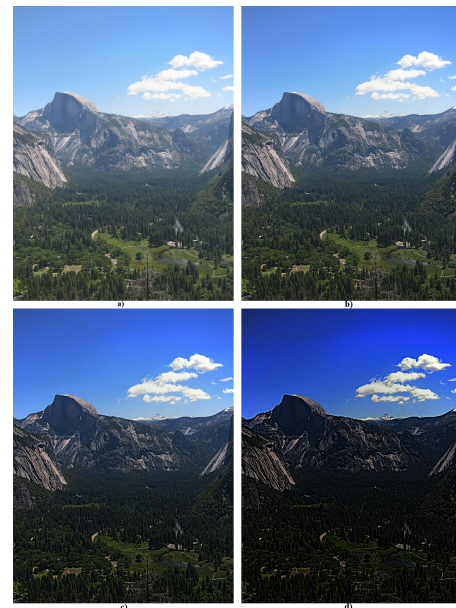


Fig. 4. Dehazed images for a) $\omega = 0.25$ b) $\omega = 0.5$, c) $\omega = 0.75$, d) $\omega = 0.9$

The transmission estimated for low ω values will result in too hazy images as seen in Fig.4.b, while high ω values will result in dark images as seen in Fig.4.d. Since, the transmission estimations obtained by the proposed prior is already similar to the image itself, no additional computation is needed to recover the dehazed image.

After observing different kinds of hazy images, ω value should be chosen around 0.75.

C. Reconstruction of the Haze Free Image

The haze free image can be recovered as:

$$J_m(x) = \frac{I_m(x) - A_m}{t_m(x)} + A_m \quad (17)$$

The proposed dehazing method can be summarized as follows:

- 1) The approximation and wavelet layers for the band of the hazy image are obtained by equations (11) and (12), respectively.
- 2) The multi-scale products for approximation and wavelet layers are calculated according to equations (13) and (14), respectively.
- 3) Obtained multi-scale products are summed up by (15) to obtain the proposed prior (AWP).
- 4) The airlight value for the corresponding band is obtained by the mean of the brightest 0.1% pixels of AWP.
- 5) The transmission is estimated by using the airlight and the proposed prior as in (16).
- 6) The haze free band of the input image is obtained by equation (17).
- 7) If the input image is a color image, above steps are repeated for the next color band of the image.

IV. EXPERIMENTAL RESULTS

In order to understand the performance of the proposed AWP method, comparisons with the traditional methods are made. The proposed AWP and the classical methods are applied to a variety of hazy images. The comparisons are made both visually and quantitatively.

A. Visual Comparison

The first hazy image is shown in Fig.5.a. The dehazing results for Tan [7], Tarel [10], Kopf [6], DCP [8], Fattal [9], MP [16] and the proposed AWP method are given in Fig.5.b-h, respectively. According to Fig.5.b, Tan's method distorts the color for overall image. Fig.5.c shows that Tarel's method can not keep the color information, especially at the far end. Kopf's method, as seen in Fig.5.d, has a good performance for far scene, while it has a poor performance in closer regions. As seen Fig.5.f, Fattal's method has a poor performance where haze is thicker, as seen in the middle and far areas, even though it performs good for closer areas. As seen in Fig.5.e, DCP method distorts the color at the far end, especially in sky region, even though having a good performance in other regions. MP method as shown in Fig.5.g, does not distort the color information and has a better performance than the aforementioned methods. However, a closer look demonstrates that the dehazing performance for middle area is not good enough. The proposed AWP method, as seen in Fig.5.h, has a better performance for overall image.

For further comparison, a scene from New York City that includes more details shown in Fig.6.a is used. The haze

removal results for Tan [7], Tarel [10], Kopf [6], DCP [8], Fattal [9], MP [16] and the proposed AWP method are shown in Fig.6.b-h, respectively. Tan's method has distorted the colors as expected. Tarel's method has a poor performance for the far scene, while Kopf's method has a poor performance for the closer region. Fattal's method can not dehaze the regions with thicker haze. DCP method overamplifies the sky region, while it can not amplify the right top of the scene. MP method being better than the former methods, can not dehaze the mid range of the image. The proposed AWP method has a better performance for the whole image.

In order to see the differences between the dehazing methods accurately, zoomed areas of Fig.6 is given in Fig.7. The zoomed version includes the sky region (far end) and mid range of Fig.6.

As seen in Fig.7, Tan's, DCP method and Tarel's method distort the color of the sky. Moreover, under the sky region, the visibility has not increased for He's, Fattal and MP method. The best dehazing is performed by the proposed AWP method. A closer look to mid area demonstrates that, Tan's and Tarel's method suffers from color distortion. Kopf, DCP, Fattal and MP methods can not dehaze the mid range well. The proposed method dehazes the mid range and increases the visibility better than the former methods.

B. Quantitative Comparison

In order to make an objective comparison, the proposed AWP method is compared quantitatively with state of art methods. The figures used are given in [12], which are available online. Three of the figures and true transmissions are included in this paper for comparison purposes. Fig.8 shows the original haze-free images and dehazed images obtained by, Fattal[9], DCP[8], CL[12], Catt[13], DehazeNET[15] and AWP methods, respectively. According to the Fig.8, the dehazed images obtained by DCP, CL, and the proposed AWP methods are seems to be the closest to the original ones.

In order to compare quantitatively, SSIM and RMSE indexes are evaluated between the original image and dehazed images by the methods abovementioned. In addition, the RMSE values are calculated between the true transmission and recovered transmissions of the methods. The RMSE and SSIM values calculated between haze-free and dehazed images are given in Table 1.

TABLE I
QUANTITATIVE COMPARISON BETWEEN HAZE-FREE AND DEHAZED IMAGES

Figure	Metric	Fattal[9]	DCP[8]	CL[12]	Catt[13]	DehazeNET[15]	AWP
Mansion	RMSE(%)	10.3819	9.2994	8.6000	18.6304	13.0867	9.0765
	SSIM	0.8838	0.7241	0.9013	0.3334	0.4752	0.8909
Flower	RMSE(%)	3.2497	3.9622	2.3262	15.9957	10.2755	2.8838
	SSIM	0.8973	0.8940	0.9099	0.5964	0.5629	0.9030
Couch	RMSE(%)	8.6370	7.5886	7.6351	14.4420	13.8691	7.1690
	SSIM	0.8930	0.8944	0.8949	0.6615	0.6568	0.9026

According to Table 1, the best SSIM and RMSE values for Mansion and Flower images are obtained by CL method followed by the proposed method. CL method is slightly better than the proposed AWP method for this figures. For Couch

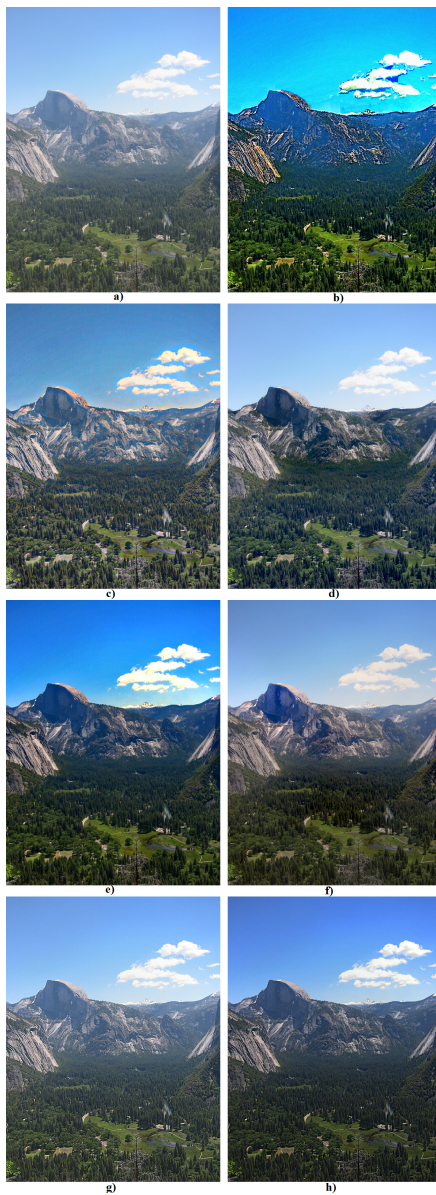


Fig. 5. a) Hazy image, Dehazing results for b) Tan c) Tarel, d) Kopf, e) DCP, f) Fattal, g) MP and h) the proposed AWP methods.

image, the best scores are obtained by the proposed AWP method.

For further comparison, recovered transmission of the methods aforementioned and the transmissions are given in Fig.9. According to Fig.9, the best recovered transmissions seems to be achieved by CL, DehazeNET and the proposed AWP methods.

The RMSE values calculated between true transmission and dehazing methods are given in Table 2.

TABLE II
RMSE VALUES BETWEEN TRUE TRANSMISSION AND RECOVERED TRANSMISSIONS

Figure	Fattal[9]	DCP[8]	CL[12]	Catt[13]	DehazeNET[15]	AWP
Mansion	25.9045	31.2422	18.0345	28.9952	17.5627	19.8263
Flower	52.7966	16.1832	12.1464	40.2297	23.9222	14.6925
Couch	27.6459	24.9502	25.6094	26.2958	25.3566	23.8337

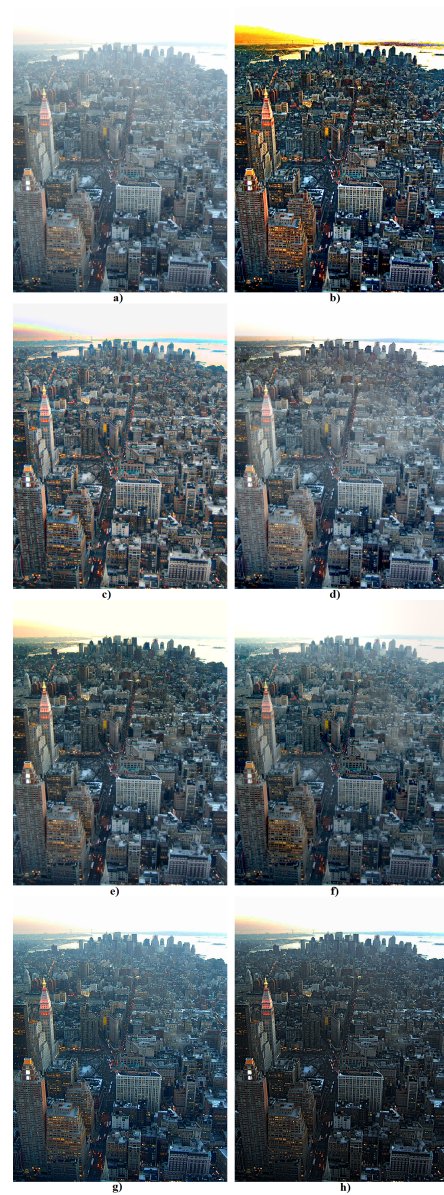


Fig. 6. a) Hazy image, Dehazing results for b) Tan c) Tarel, d) Kopf, e) DCP, f) Fattal, g) MP and h) the proposed AWP methods.

According to Table 2, the best RMSE value for Mansion image is obtained by DehazeNET method followed by the proposed AWP and CL methods. For Flower image, the best score is achieved by CL followed by the proposed method. For Couch image, the best score is obtained by the proposed AWP method.

Experiments are performed in Matlab with a 2.4 GHz Intel i3 processor PC. The average running time of the methods are tabulated in Table 3. It can be seen that the proposed AWP method is nearly 30 times faster than the DCP method and 10 times faster than Catt method.

Taking into consideration both visual and quantitative comparisons, the proposed AWP method is either superior to the classical methods or in good agreement with the results of the state-of-art methods.

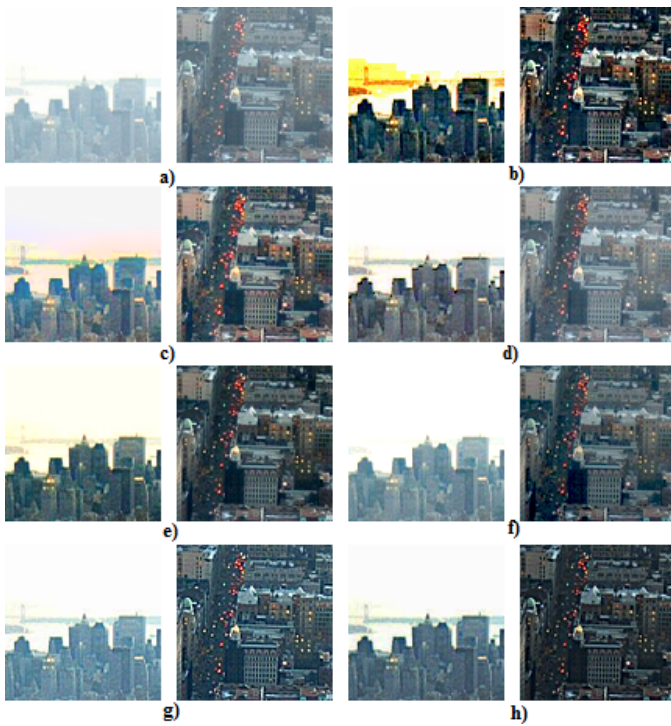


Fig. 7. A closer look to Fig.6. a) Hazy image, Dehazing results for b) Tan c) Tarel, d) Kopf, e) DCP, f) Fattal, g) MP and h) the proposed AWP methods.

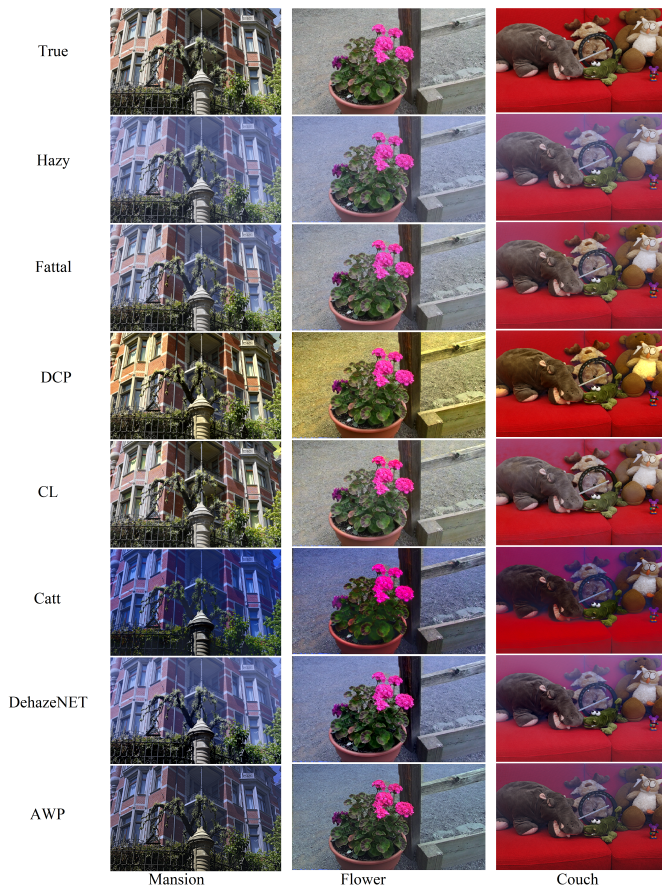


Fig. 8. Comparison with synthetic hazy images.

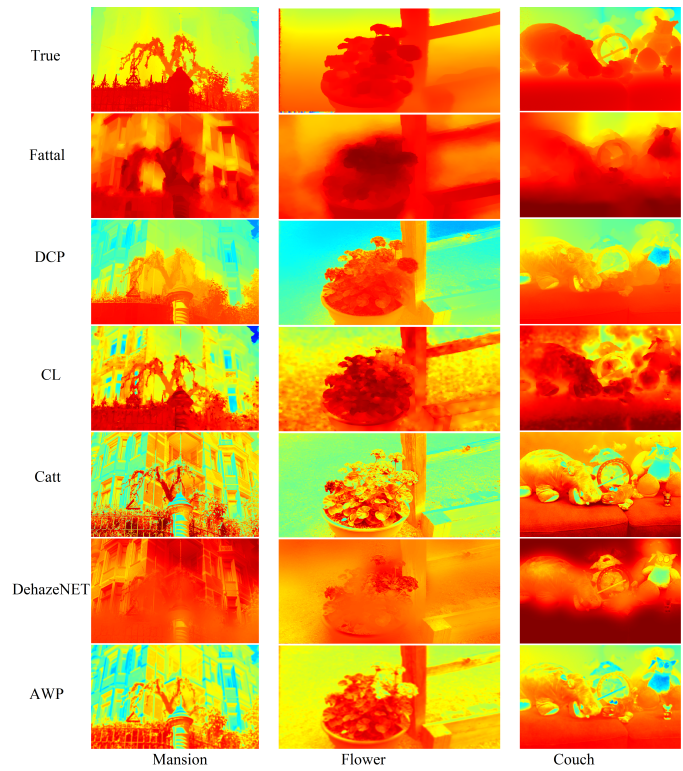


Fig. 9. Comparison of recovered transmissions with synthetic hazy images.

TABLE III
AVERAGE TIME CONSUMPTION OF THE METHODS FOR 1024X640 SIZED IMAGES

Method	Fattal[9]	DCP[8]	Catt[13]	DehazeNET[15]	AWP
Time Consumption (sec)	4.41	9.98	3.52	6.73	0.32

V. CONCLUSION

In this paper, a novel prior for single image dehazing is proposed. In order to obtain the proposed prior, first additive wavelet transform is applied to the hazy image to obtain the approximation and wavelet layers. The adjacent approximation and wavelet layers are multiplied and summed up to calculate the proposed prior. This new prior detects the haze present in the image, as well as it keeps the important details of the image. Using the proposed prior alongwith the commonly used hazy image model, an effective dehazing method is obtained. The visual and quantitative comparisons with traditional methods demonstrate that the proposed method outperforms the traditional methods.

REFERENCES

- [1] T. K. Kim, J. K. Paik, and B. S. Kan, "Contrast enhancement system using spatially adaptive histogram equalization with temporal filtering," *IEEE Transactions on Consumer Electronics*, vol. 44, no. 1, pp. 82–87, 1998.
- [2] J. A. Stark, "Adaptive image contrast enhancement using generalizations of histogram equalization," *IEEE Transactions on Image Processing*, vol. 9, no. 5, pp. 889–896, 2000.
- [3] J.-Y. Kim, L.-S. Kim, and S.-H. Hwang, "An advanced contrast enhancement using partially overlapped sub-block histogram equalization," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 11, no. 4, pp. 475–484, 2001.

- [4] S. G. Narasimhan and S. K. Nayar, "Contrast restoration of weather degraded images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 6, pp. 713–724, 2003.
- [5] S. Shwartz, E. Namer, and Y. Y. Schechner, "Blind haze separation," in *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06)*, vol. 2, 2006, pp. 1984–1991.
- [6] J. Kopf, B. Neubert, B. Chen, M. Cohen, D. Cohen-Or, O. Deussen, M. Uyttendaele, and D. Lischinski, "Deep photo: Model-based photograph enhancement and viewing," *ACM Trans. Graph.*, vol. 27, no. 5, pp. 116:1–116:10, 2008.
- [7] R. T. Tan, "Visibility in bad weather from a single image," in *2008 IEEE Conference on Computer Vision and Pattern Recognition*, 2008, pp. 1–8.
- [8] K. He, J. Sun, and X. Tang, "Single image haze removal using dark channel prior," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 12, pp. 2341–2353, 2011.
- [9] R. Fattal, "Single image dehazing," *ACM Trans. Graph.*, vol. 27, no. 3, pp. 72:1–72:9, 2008.
- [10] J. P. Tarel and N. Hautière, "Fast visibility restoration from a single color or gray level image," in *2009 IEEE 12th International Conference on Computer Vision*, 2009, pp. 2201–2208.
- [11] G. Meng, Y. Wang, J. Duan, S. Xiang, and C. Pan, "Efficient image dehazing with boundary constraint and contextual regularization," in *2013 IEEE International Conference on Computer Vision*, 2013, pp. 617–624.
- [12] R. Fattal, "Dehazing using color-lines," *ACM Trans. Graph.*, vol. 34, no. 1, pp. 13:1–13:14, 2014.
- [13] Q. Zhu, J. Mai, and L. Shao, "A fast single image haze removal algorithm using color attenuation prior," *IEEE Transactions on Image Processing*, vol. 24, no. 11, pp. 3522–3533, 2015.
- [14] I. Riaz, X. Fan, and H. Shin, "Single image dehazing with bright object handling," *IET Computer Vision*, vol. 10, no. 8, pp. 817–827, 2016.
- [15] B. Cai, X. Xu, K. Jia, C. Qing, and D. Tao, "Dehazenet: An end-to-end system for single image haze removal," *IEEE Transactions on Image Processing*, vol. 25, no. 11, pp. 5187–5198, 2016.
- [16] N. H. Kaplan, K. K. Ayten, and A. Dumlu, "Single image dehazing based on multiscale product prior and application to vision control," *Signal, Image and Video Processing*, 2017.
- [17] C. Qin and X. Gu, "A single image dehazing method based on decomposition strategy," *Journal of Systems Engineering and Electronics*, vol. 33, no. 2, pp. 279–293, 2022.
- [18] Z. Zhu, H. Wei, G. Hu, Y. Li, G. Qi, and N. Mazur, "A novel fast single image dehazing algorithm based on artificial multiexposure image fusion," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1–23, 2021.
- [19] H. Ullah, K. Muhammad, M. Irfan, S. Anwar, M. Sajjad, A. S. Imran, and V. H. C. de Albuquerque, "Light-dehazenet: A novel lightweight cnn architecture for single image dehazing," *IEEE Transactions on Image Processing*, vol. 30, pp. 8968–8982, 2021.
- [20] Y. Huang and X. Chen, "Single remote sensing image dehazing using a dual-step cascaded residual dense network," in *2021 IEEE International Conference on Image Processing (ICIP)*, 2021, pp. 3852–3856.
- [21] J. Nunez, X. Otazu, O. Fors, A. Prades, V. Pala, and R. Arbiol, "Multiresolution-based image fusion with additive wavelet decomposition," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 37, no. 3, pp. 1204–1211, 1999.
- [22] K. He, J. Sun, and X. Tang, "Guided image filtering," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 6, pp. 1397–1409, 2013.
- [23] S. G. Narasimhan and S. K. Nayar, "Vision and the atmosphere," *International Journal of Computer Vision*, vol. 48, no. 3, pp. 233–254, 2002.
- [24] Y. Chai, H. Li, and M. Guo, "Multifocus image fusion scheme based on features of multiscale products and pcnn in lifting stationary wavelet domain," *Optics Communications*, vol. 284, no. 5, pp. 1146 – 1158, 2011.

processing. His primary research interests include digital signal and image processing.

Nur Huseyin Kaplan received his B. Sc., M. Sc, and Ph.D degrees in Electronics and Telecommunication Engineering from Istanbul Technical University, Turkey. He worked for TURKSAT A.Ş. as a specialist between 2005-2015. Since 2018, he is an Associate Professor at Electrical and Electronics Engineering Department with Erzurum Technical University, Turkey, where he teaches undergraduate/graduate level courses about signal and image

Classification Of Hand Images by Person, Age and Gender with The Median Robust Extended Local Binary Model

Emrah Aydemir and R.T. Esfandiyar Alalawi


Abstract— Biometric technologies try to automatically recognize individuals by considering the physiological and behavioral characteristics of individuals. Although the methods used here are very diverse, the personal qualities used also vary. Facial features, finger and vein prints, iris, retina, ear, hand, and finger recognition are only some of the physiological features. It may be preferred to use one or more of these personal features to reduce the margin of error that may arise depending on the security level in the applications used. Biometric recognition systems have varying requirements in security systems applications. Fingerprint and iris recognition work well in applications that require high security levels, while applications that require low security levels are not suitable due to privacy concerns. On the other hand, identification from hand images is more accepted based on the idea that it does not have a very high distinctiveness. But it is sufficient for medium security applications. Apart from these, palm images have many advantages such as reliability, stability, user-friendliness, non-intrusiveness, and flexible use. In this study, it is aimed to identify people, determine their ages, and determine their gender by using both upper surface and inner surface images of right-left hand data of hand shape. For this purpose, images of both the inner surface of the hand (10) and the outer surface of the hand (10) of 100 different people were collected. This was done separately for the right and left hands, and a total of 3955 images were obtained. The features of these images were extracted using the Median Robust Extended Local Binary Model (MRELBP). Images are classified for person, age and gender. The results were 91.4%, 85.9% and 92.6%, respectively.

Index Terms— Hand image, Median Robust Extended Local Binary Model, MRELBP, Classification.


I. INTRODUCTION

BIOMETRIC IDENTIFICATION has become inevitable due to rapid developments in the world and increasing security requirements.

EMRAH AYDEMİR, is with Department of Electrical Engineering University of Sakarya University, Sakarya, Turkey, (e-mail: emrahaydemir@sakarya.edu.tr).

 <https://orcid.org/0000-0002-6598-7251>

RAGHAD TOHMAS ESFANDIYAR ALALAWI, is with Department of Computer Engineering, Ahi Evran Uni., Kirsehir, Turkey, (e-mail: alalawiraghad@yahoo.com).

 <https://orcid.org/0000-0002-6598-7251>

Manuscript received September 6, 2022; accepted November 22, 2022.

DOI: [10.17694/bajece.1171905](https://doi.org/10.17694/bajece.1171905)

Now, the sustainability of security is low with the records kept in the books or traditional applications such as photographs, identity cards and signatures. It has become a necessity to carry out these processes with computer-based systems that will reduce human error. In addition to keeping records in various databases, digitally obtained biometric data should be used for person verification. These systems will be able to see the small details and differences that the human eye misses much better. For this, they match the data stored in the databases. Such systems have become frequently encountered in daily life. In many countries, palm, eye, fingerprint, etc. automatically at airports and border gates. biometric recognition systems are used. Similarly, in many countries, studies are carried out to record biometric data in passports [1].

There are many advantages to using biometric data to identify people. The most important of these are the uniqueness of biometric data and having a fixed shape. As the demand for biometric recognition systems increases, the demand for the evaluation of their performance also increases accordingly. However, not all biometric recognition systems have the same advantages. For the most widely used fingerprint recognition systems, the finger is expected to be clean and free of various stains. Such situations negatively affect performance [2]. Interest in ear recognition studies and their conversion to commercial technologies is low for several reasons [3, 4]. Like many other biometric features, palm shape changes little with age. In addition, the veins on and outside the palm are also used as a unique element in person recognition [5]. Because the vein patterns are based on the body's internal biological information, they are not affected by the destructions of the outer surface and cannot be easily copied. This increases its reliability as a biometric recognition feature. The appearance of these checkers in different individuals is a distinctive feature [6].

The palm consists of the inner surface of the hand between the wrist and the fingers. This inner surface consists of three parts: finger root region, inner region and outer region. There are three main lines on the palm that are formed because of flexing the hand. These are the life line, the heart line, and the head line [7]. These lines and accompanying wrinkles and secondary lines become more distinctive over time. It characterizes people because of their stable nature. An example palm image is given in the figure below.

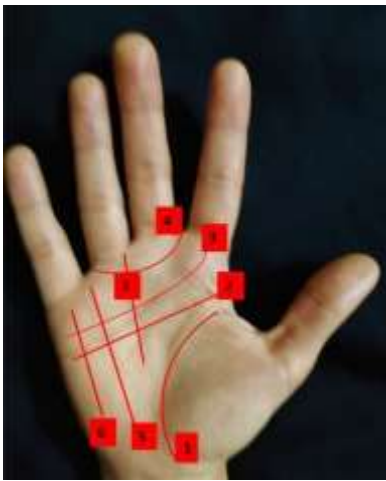


Fig. 1. An example palm shape and lines

The number of lines on the palms, their relationships and distances are the main features used in authentication. Figure 1 above shows both the three main lines and the crumpled lines. The names of the lines given with numbers on the figure are as follows.

1. Lifeline
2. Headline
3. Heart Line
4. Belt of Venus
5. Sun Line
6. Mercury Line
7. Destiny Line

In addition to being authentic, they are practical for authentication, as the lines on the palm are more pronounced rough lines. It offers the opportunity to work quickly and flexibly even with lower resolution images [8]. Systems that extract features from palm images by focusing only on line structures have certain disadvantages. Some of these are the indistinct palm prints, palm prints with similar line features, and ignoring the thickness and width of the lines [9]. Wrinkles are fine and ridges can be analyzed with a medium resolution, although low resolution is not a problem because the main lines in the palm images are thick.

Chiromancy is the study of the size, shape and appearance of the hand, including analysis of the color, texture and elasticity of the skin. Chiromancy studies begin with recognizing the basic hand shapes. Evaluation of hand shape gives an idea about the basic character of the individual, and throughout history, palmists have developed various classification systems for hand types [10]. Saint -Germain [11] classified basic personalities according to hand types. Finger proportions, hand lines and other factors are considered here.

Knowing the shapes of people's hands and palms means that these processes can be easily determined by computer-based systems. Fingers and fingerprints are also involved in the recognition systems of the inner and outer surfaces of the hand. Fingerprints are caused by the different growth and pressure experiences of the embryo inside the womb. Therefore, it differs even for identical twins [12, 13]. Folds, angles and distances in fingerprints are used for identification [14]. But for this, the image must be clear enough to distinguish these subtleties.

A. Aim

Biometric technologies try to automatically recognize individuals by considering the physiological and behavioral characteristics of individuals. Although the methods used here are very diverse, the personal qualities used also vary. Facial features, finger and vein prints, iris, retina, ear, hand and finger recognition are only some of the physiological features. It may be preferred to use one or more of these personal characteristics in order to reduce the margin of error that may arise according to the security level in the applications used [15]. Biometric recognition systems have varying requirements in security systems applications. Fingerprint and iris recognition work well in applications that require high security levels, while in applications that require low security levels, they are not suitable due to privacy concerns. On the other hand, identification from hand images is more accepted based on the idea that it does not have a very high distinctiveness. But it is sufficient for moderate security applications [16]. Apart from these, palm images have many advantages such as reliability, stability, user-friendliness, non-intrusiveness and flexible use. In this study, it is aimed to identify people, determine their ages and determine their gender by using both the upper surface and inner surface images of the right-left hand data of the hand shape.

B. Literature review

Images from palm prints are available from Fei et al. [17] using the multi-layered direction subtraction method. With this method, firstly, the apparent direction on the surface layer of the palm print is obtained, and then the hidden features in the energy map layer of the visible direction are used. The histogram of both features is extracted and PolyU 98.83% recognition success is achieved with the dataset. PolyU with the same dataset multispectral palm print images Cui [18] with the fusion strategy. Here, images are captured in red, green, blue and near infrared, respectively. Thus, it contains more information compared to other images. It used the expanded general color image discriminant (GCID) model to generate three new color components to further improve the recognition performance, achieving 99.57% success. Zhang and Gu [19] proposes to combine competitive coding method and two-step test sample sparse representation (TPTSR) method for palmprint recognition. The TPTSR method, which is one of the representative methods, takes the entire palm image as input and determines the contribution of the training samples of each class in representing the test sample. TPTSR also uses the corresponding contribution values to calculate the similarities between the test sample and each class. The competitive coding method is a feature-based method and is highly complementary to TPTSR. The author uses a weighted fusion scheme to combine the matching scores generated from the TPTSR and competitive coding method. Experimental results show that the proposed method can achieve a very high classification accuracy and outperform both TPTSR and competitive coding method. The method, validated with 6000 samples from 500 people (12 palms), and the match rate of TPTSR is presented, reveals a 94% success rate. Zhang and Gui [20] propose a palm recognition method based on representation in feature space in another study. The proposed method aims to represent the test

sample as a linear combination of all training samples in the feature space and then uses the resulting linear combination to perform palmprint recognition. The author applies the original space-to-feature space mapping using kernel functions such as the radial basis function (RBF). In this method, the selection of the kernel function's parameter is important. It proposes an automatic algorithm for selecting the parameter. The basic idea of the algorithm is to optimize the feature space such that while instances of the same class are clustered well, instances from different classes are pushed away. The proposed criterion measures the goodness of a feature space, and by minimizing this criterion, the optimal kernel parameter is obtained. Experimental results on the database achieved 98% success in the public dataset of the proposed methods.

Li, Cao and Lu [21] that the two-stage test sample representation method (TPTSR) performs very well. TPTSR is not only a competent representational classification method, but also computationally more efficient than the original sparse representation methods. However, although TPTSR seems suitable for palm recognition, it has not been widely tested and it is unknown how to set the parameter (number of nearest neighbors) which is absolutely crucial for real-world applications. The study aimed to analyze the performance of the method in palm identification. In this context, numerous experiments on palm recognition have been conducted to address the mentioned problems. The method was validated on the "PolyU" database and achieved low error rates. Zhao and Zhang [22] propose a general framework that represents high-level discriminating features for multiple scenarios in palmprint recognition with learned discriminating deep convolutional networks called deep discriminant representation (DDR). They use discriminant deep convolutional networks for learning, with limited palmprint training data used to extract deep discriminants. Next, the collaborative representation-based classifier for palmprint recognition is applied, which is flexible and practical in multiple scenarios. Experimental results show that DDR produces the best recognition performance in general palm recognition compared to other state-of-the-art methods. Over 98% results were obtained in the PolyU Multi-spectral database with DDR, M_R, M_B, M_G and M_NIR respectively for contact-based palm recognition under different lighting sources, and in DDR, IITD and CASIA databases for contactless palmprint recognition.

Matkowski, Chai, and Kong [23] to examine palmprint identification on images collected in an uncontrolled and uncooperative environment. It contains 2035 palms collected from the internet and obtained from 7881 images. The proposed algorithm consists of an alignment network and a feature extraction network and can be trained end-to-end. The algorithm was compared with state-of-the-art online palm recognition methods and evaluated on three common contactless palmprint databases, IITD, CASIA and PolyU, and two new databases, NTU-PI-v1 and NTU contactless palmprint database. The proposed algorithm achieved successful results in NTU-PI-v1, NTU-CP-v1, IITD, PolyU and CASIA databases at a rate of 41.92%, 95.34%, 99.61%, 99.77% and 97.65%, respectively. This system is sufficiently resistant to

noise and deformation [24]. In the identification system, ONPP is used for manifold learning and the point-to-manifold distance function is used to test the finger core. Based on the infrared image database TED-FV, the proposed method achieves 97.8% recognition rate with the identification pattern, which is better than other traditional methods. Experimental results have shown that the proposed system is an effective and robust method for its performance.

Hong et al. [25] multispectral palmprint instead of natural light palm print and developing a multi-band palm print recognition method based on a hierarchical idea, to achieve higher recognition rate with more distinctive information. First, Block-Based Routing Code (BDOC) as a rough feature, and Block-Based Directed Gradient as a good feature. They extract the histogram (BHOG). Secondly, a hierarchical recognition approach is proposed based on these two types of features. They combine different features from different bands in the proposed scheme to improve recognition accuracy. Finally, the experimental results show that the recognition accuracy of the proposed method is not only superior to previous high-performance methods based on the PolyU palm database with natural light but can further improve the state-of-the-art performance achieved by some approaches. Trying to refine coding-based approaches using multi-spectral palm print images, Bounanche et al. [26] presents a new multi-spectral palmprint recognition approach based on directed multi-scale log-Gabor filters. The proposed method aims to increase recognition performances by suggesting new solutions in three stages of the recognition process. Inspired by bitwise competitive coding, feature extraction uses multi-resolution log-Gabor filtering; where the final feature map consists of the winning codes of the bank response of the lowest filters. MS-PolyU EER results of less than 0.01 were obtained in the database. Ma et al. [27] introducing the idea of discrimination analysis into palm coding, they attempted to extract DOSFL, orientation and scale features with more appropriate discrimination. DOSFL then uses four code bits to represent both the orientation and scale characteristics of the palmprint and uses the Hamming distance for code matching. In order to make better use of the direction and scale information contained in the palm print samples, a discriminative learning (MOSDL) approach with versatile and multi-scale features for palm print recognition, which can effectively combine different orientation and scale feature data in the discriminative learning process, is proposed in this context. Experimental results on two publicly available palm databases, including the HK PolyU database and the UST image database, show that the proposed method can achieve better recognition results than the compared methods, with accuracy values of 98.05% and 97.63%, respectively.

Tamrakar et al. [28] presented an effective palm print recognition technique for palm prints collected with both visible and multispectral imaging systems. ROI extraction is a challenging task for palm print captured in unconstrained environment. ROI achieved by finger spacing and palm width makes system rotation and translation invariant. The

approximate ROI obtained by first-order decomposition of the ROI reduces the Haar wavelet computational overhead and noise. Phase quantization with Gaussian derivative filter of AROI gives the Gaussian derivative phase pattern image, and its block-based histograms are combined to form a single vector called the BGDPPH descriptor. Size reduction is achieved by increasing the distinction between genuine and dishonest scores using Chi -RBF kernel discriminant analysis (KDA). Weighted score level fusion of spectral palm prints in the Fisher criterion improves the recognition rate. The robustness of the proposed BGDPPH identifier against blur and noise has been validated on four grayscale and two multi-band palmprint databases collected via touch-based devices. It is stated in the study that when the proposed method is applied on these four different data sets, "PolyU 2D", "IITDMJ", "CASIA", and "IITD" accuracy rates of 99.98%, 100%, 99.22%, and 99.19%, respectively. Hong et al. [29] presented a palm print recognition system using the fast Vese-Osher decomposition model to process blurry palm print images. First, a Gaussian blur distortion model (GDDM) is proposed to characterize image blur, and from this model it is observed that there are some stable features in palm images with different blur scale. Second, the structure layer and texture layer of blurred palm print images are obtained using the fast Vese-Osher decomposition model, and the structure layer proves to be more stable and robust than the texture layer for palm recognition. As a result, a new algorithm based on the weighted histogram of the directed gradient for the locally selected model (WHOG-LSP) is proposed and used to extract some robust features from the structure layer of the blurred palm print images. These extracted features were used to fix poor performance issues associated with translation and rotation in palm print recognition. Finally, the normalized correlation coefficient (NCC) was used to measure the similarity of palm features for the proposed recognition system. Extensive experiments on the PolyU palm database and the fuzzy PolyU palm database confirm the effectiveness and real-time feasibility of the proposed recognition system. Additionally, IIT Delhi Touchless An experiment was conducted to verify the robustness of the proposed method in the Palmprint database and EER values of 0.0727 % and 0.9210 were obtained, respectively.

II. MATERIAL AND METHOD

A. Data Gathering

Before starting the study, the people who will collect image data were informed about the purpose of the study and how the data will be used. It was stated that there was voluntary participation, and it was clearly explained that no data that could identify him, such as his name or surname, would be collected. Voluntary participants were asked to press the inner and outer surfaces of each hand on a piece of black cloth. In the study, image data of both the palm and the upper surface of the hand of 100 different people were collected. A total of 40 images were obtained from a person, 10 for the left and right palms, and 10 for the upper surface of the hand. In this way, 3955 images were obtained. In the images that were examined

later, it was seen that 5 files had the same name, and they were deleted because it was not known exactly to which user they belonged. For this reason, the number of files that should have been 4000 in total became 3955. Images were obtained from different angles with a high-resolution device. Only age and gender information were obtained from the people whose images were taken. Apart from this, no data to identify individuals was recorded. In order to determine who the people are, a sequential ID number in the range of 1-100 is given. A naming convention has been developed for the name of each image file so that there are many images and no name confusion. According to this rule, people's ID numbers, ages, genders and sequence number of the captured image are used. The underscore (_) character is used to legibly parse each data. Here, the following rules have been applied respectively.

- **Person ID:** It is a unique code value for each person photographed. This value ranges from 1 to 100.
- **Age:** The age is written directly as a number to express how old the person is.
- **Gender ID:** It is expressed with the value of 1 if the person photographed is male, and 0 if it is a female.
- **Photo ID:** Since more than one image was taken of each person, each image was numbered sequentially from 1 to 40.

The inner and outer surface images of the left hand of the person with the ID number 49 are shown in the figure below.



Fig. 2. Inner and outer space surface view of a sample person's hand

Naming the files according to the naming convention both prevents confusion and makes it easier to determine which class they belong to during the analysis. In the figure below, some of the image files are shown in the folder.



Fig. 3. folder view of the captured image files

For the data collected in the study to be used by other researchers, all image files were uploaded to the KAGGLE dataset repository. This dataset can be downloaded from www.kaggle.com/dataset/c1b48bedfa86c1b51e6acafe8c75e08

c01dbdf7d2444a07428540ece310b8163 and used by citing this study.

B. Analysis of Data

In order to make simple statistical analyzes from the data collected in the study, first of all, the file names were transferred into an excel file. Thus, each data name in the filenames has been split into columns. Python program was used for these operations. In the analysis of the data, a computer with Windows operating system, 40GB RAM and Intel(R) Core (TM) i7-8565U CPU, 1.80GHz 1.99GHz processor and 1TB SSD was used. First of all, simple statistical findings belonging to person, gender and age categories were obtained over the file names. Thus, the distribution of the data was examined. Then, the Median Robust Extended Local Binary Model (Median) is used for feature extraction from the image files. robust extended local binary Pattern -MRELBP) method was used. Local Binary Pattern (LBP), Local Phase Quantization (LPQ), Local Ternary Pattern (LTP) and MRELBP feature extraction methods are used with Subspace KNN classification algorithm. Success rates in person classification were found to be 75.1, 76.2, 61.9, and 79.3, respectively. This method was used because the highest accuracy rate was obtained in the MRELBP method. This method [30] as a tool in the Matlab program. Thus, 1x800 vectors were produced for each image file. This process has been done for all files. First, to identify people from hand images, person ID values were written as classification tags in these files. Then, another attribute file was created, and gender was written as the class label, and another file was created, and age was written as the class label. Thus, three different attribute files were obtained. Images are not separated as the inner surface or outer surface of the hand, nor are they separated as right and left hands. The data were separated into training and testing using the 10-fold cross-validation method. Then, these feature files were tested with 32 different classification algorithms in the Matlab program. The highest classification success was tried to be achieved. For classification processes, only one learning and testing process was performed. Repeated attempts were not made to achieve a higher success rate. In the title of Findings, each classification algorithm and the success rate obtained are given.

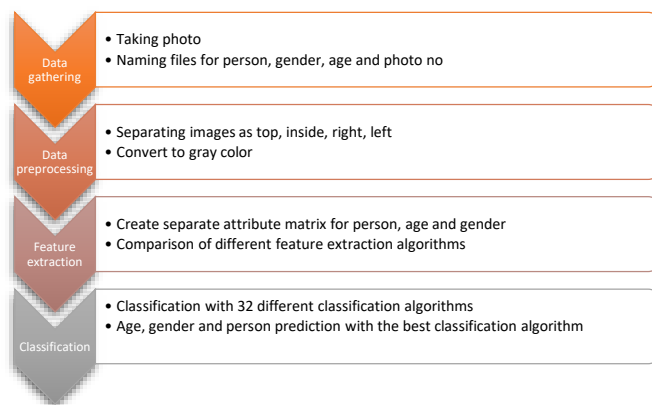


Fig. 4. Operation method and steps

The classification algorithms used are listed below. Default parameters are used by Matlab for these algorithms.

- Fine Tree

- Medium Tree
- Coarse Tree
- Linear Discriminant
- Quadratic Discriminant
- Logistics Regression
- Gaussian Naive Bayesian
- Kernel Naive Bayesian
- Linear Svm
- Quadratic Svm
- Cubic Svm
- Fine Gaussian Svm
- Medium Gaussian Svm
- Coarse Gaussian Svm
- Fine Knn
- Medium Knn
- Coarse Knn
- Cosine Knn
- Cubic Knn
- Weighted Knn
- Boosted Trees
- Bagged Trees
- Subspace Discriminant
- Subspace Knn
- Rusboosted Trees
- Narrows Neural Network
- Medium Neural Network
- Wide Neural Network
- Bilayered Neural Network
- Tripping Neural Network
- Svm Kernel
- Logistics Regression Kernel

C. Feature Extraction with Median Robust Extended Local Binary Model Extraction of Median robust extended local binary Pattern -MRELBP)

Local binary patterns (LBP) are an image processing algorithm that extracts texture features from images and uses them to classify and segment images. This algorithm was first proposed by computer scientists Oyala and Pettenkainen in 1994. Since then, this algorithm has evolved to include: It was used to extract the Tissue feature from 3D images and video clips and to recognize faces in medical images and identify diseases. Local binary patterns are calculated for a texture image in its simplest form as follows:

- First of all, the neighborhood system next to each pixel is used. That is, 8 values around the midpoint of a 3x3 matrix are taken into account.
- Each of the adjacent points (eight points) is compared to the center point. The value “0” is stored if the center point is larger, otherwise “1” values are retained.
- Finally, these 1 and 0 values obtained for each pixel in the image are converted to a decimal number. Thus, the LBP code consisting of decimal numbers is generated for each pixel.
- A histogram of all decimal values in the image is created, which is then used as a description of the entire image texture.

An example of a pixel for the above-mentioned calculation is given with steps in the figure below.

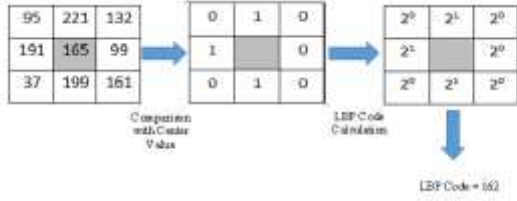


Fig. 5. LBP Sample Calculation

The following mathematical formulas are used to calculate the LBP code.

$$LBP_{8,1}(x_c) = f(x) = a_0 + \sum_{p=0}^{p-1} u(u(x_p - x_c))^{2^p} \quad (1)$$

$$u(y) = \begin{cases} 1 & y \geq 0 \\ 0 & y < 0 \end{cases} \quad (2)$$

In these equations, it is used to express the difference between the neighboring pixel and the center pixel y , the LBP tag represents the generated center pixel x_c , it x_p is expressed by the neighbors of the center pixel, $u(y)$ the result is the result, and here the result is the result of the LBP operator [31].

The upper version of LBP, ELBP, is designed to calculate distinctive spatial relationships in a local region. LBP, on the other hand, only calculates the relationship between neighbors and a central point. For these reasons, the ELBP method contains more spatial information than LBP. In addition, ELBP consists of three parts. With ELBP_CI, the central pixel is taken and the relationship with the neighbors is calculated accordingly. The following equation is used for this calculation.

$$ELBP - CI_{r,p}(x_c) = \sum_{n=1}^{\infty} s(x_c - \beta) \quad (3)$$

In ELBP_NI, on the other hand, instead of calculating the central pixel, the average of the neighbors is calculated, and this creates a serious functional difference from the LBP. The ELBP_NI formula is given in Equation 4.

$$ELBP - NI_{r,p}(x_c) = a_0 + \sum_{n=0}^{p-1} s(x_{r,n} - \beta_{r,p})^{2^n} \quad (4)$$

Alongside the ELBP_CI and ELBP - NI values, the ELBP_RD value is also calculated. The ELBP_RD equation 2.5 derived from the pixel differences in the radial directions is calculated.

$$ELBP - RD_{r,r-1,p}(x_c) = \sum_{n=0}^{p-1} s(x_{r,p,n} - x_{r-1,p,n})^{2^n} \quad (5)$$

One of the ELBP's drawbacks is that it is very inefficient against image noise. That's why MRELBP was developed. The solution is to first replace the individual pixel intensities at a point with representation over a region. Also, a single pixel that is not very bright or has distorted noise can significantly affect the Gaussian smoothed value or average value of all neighboring pixels, resulting in unreliability of the binary code. For these reasons, MRELBP was developed. The created MRELBP consists of three parts MRELBP CI, MRELBP NI

and MRELBP RD.

$$MRELBP_{CI}(x_c) = s(\phi(X_{c,w}) - \mu_w) \quad (6)$$

$X_{c,w}$ Represents the central, $\phi(X_{c,w})$ average value for the area that specifies the x_c local patch. $w * w$

$$MRELBP_{NI_{r,p}}(x_c) = \sum_{n=0}^{p-1} s((\phi(X_{r,p,w_r,n}) - \mu_{r,p,w_r,n}))^{2^n} \quad (7)$$

$$\mu_{r,p,w_r} = \frac{1}{p} \sum_{n=0}^{p-1} \phi(X_{r,p,w_r,n}) \quad (8)$$

$X_{r,p,w_r,n}$ Specifies the $w * w$ local patch for area r, p, w_r, n centered pixels for neighbors.

$$MRELBP_{RD_{r,r-1,p,w_r,w_r-1}} = \sum_{n=0}^{p-1} s((\phi(X_{r,p,w_r,n}) - \phi(X_{r-1,p,w_r-1,n})))^{2^n} \quad (9)$$

MRELBP is very efficient and achieves very good results when compared to RELBP. The MRELBP has been extensively tested with different types of noise, such as Gaussian white noise, Gaussian blur, salt, pepper and pixel random distortion, and delivers very efficient results. Also, MRELBP has multiple very important features, and foremost among these is computational simplicity, as well as nice features like gray scale and rotation invariance. Finally, MRELBP shows efficient results very easily without pre-training and adjusting parameters, which means convenience for the user [32].

D. Success Criteria

Different criteria can be used to evaluate the success of machine learning algorithms. But commonly used are the correct classification rate and confusion matrix. The number of correct predictions divided by the total number of samples will give the correct classification rate of the algorithm. This criterion will provide a correct interpretation when the samples belonging to each class are equal. It is calculated with the following formula.

$$\text{Doğru sınıflandırma oranı} = \frac{\text{Doğru sınıflandırılan örnek sayısı}}{\text{Toplam örnek sayısı}}$$

The confusion matrix is used to examine the classification success in more detail. This criterion indicates the full performance of the model. It can also be called an error matrix. The confusion matrix for a binary classification would be as follows.

TABLE I
CONFUSION MATRIX

	Prediction Values	
	TN	FP
Actual Values	FN	TP

The abbreviations in this matrix will be as follows.

- True Positives (TP): These are examples where the true value is 1 and the predicted value is 1. So it was predicted correctly.
- True Negatives (TN): These are instances where the true value is 0 and the predicted value is 0. So it was predicted correctly.
- False Positives (FP): These are instances where the predicted value is 1 while the true value is 0. So it was guessed wrong.

- False Negatives (FN): These are instances where the predicted value is 0 while the true value is 1. So it was guessed wrong.

Many different success measures can be calculated from the values on this matrix. The aim here is to increase the values in the correctly guessed cross column and decrease the values outside it.

III. RESULTS

A. Statistical Findings

File names were used to make various statistical calculations according to the collected person information of the data. Each data in the image filenames was parsed and subjected to simple statistical analysis. When the image data obtained in the study is examined according to gender values, it is seen that the majority of them are women. The whole sample consists of 1597 male (40%) and 2358 female (60%) data. This is illustrated visually in the figure below.

TABLE II
DISTRIBUTION OF DATA BY GENDER AND AGE

Gender	Person Number	Age Min-Max (Avg±Std.)
Male	40	10- 68 (31.12 ± 15.85)
Female	59	10- 71 (22.96 ± 12.08)
Total	99	10-71 (26.26 ± 14.23)

When the distribution of the data by age is examined, it is seen that there are different numbers of people in this range, the youngest being 10 years old and the oldest 71 years old, and it is predominantly under 25 years old. In the figure below, the distribution of the number of people of each age is given graphically. Here, it is seen that the data is especially high in the 10-20 age range and the density is very low in the 40-72 age range. It should be noted that the data here is not the number of people, but the number of images obtained from each person.

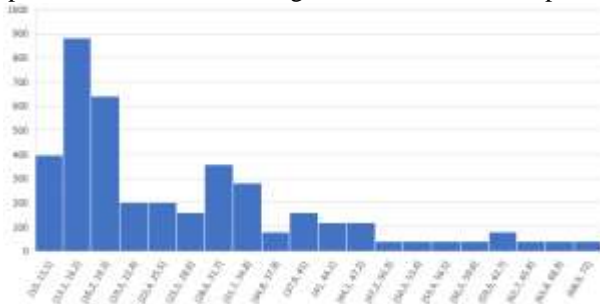


Fig. 6. Graph showing how many people are of each age

B. Person Classification Findings

There are different numbers of image data belonging to a total of 100 people. All these data are feature extracted with MRELBP. Classification algorithms were used to determine who among the 100 people on these images were. The classification rates obtained by using 32 different classification algorithms in the Matlab program and the names are given in the table below. Subspace has the highest classification success compared to other classification algorithms. It was obtained as 91.4% with the discriminant algorithm. Among other algorithms other than this, there are algorithms that find success

below or above 80%. Here, some algorithms could not produce results and therefore a total of 28 algorithms were used. Classification processes produce different results each time they run, but for this study, each algorithm was run only once.

TABLE III
PERSON CLASSIFICATION RESULTS

Algorithm Group	Algorithm Name	Accuracy rate (%)
Decision Trees	Fine Tree	38.4
	Medium Tree	15.5
	Coarse Tree	4.9
Naive Bayesian	Kernel Naive Bayesian	58.9
SVM	Linear SVM	83.5
	Quadratic SVM	86.9
	Cubic SVM	85.7
	Fine Gaussian SVM	16.9
	Medium Gaussian SVM	76.3
	Coarse Gaussian SVM	64.5
Nearest Neighbor	Fine KNN	82.8
	Medium KNN	77.5
	Coarse KNN	51.4
	Cosine KNN	74.8
	Cubic KNN	64.4
	Weighted KNN	81.6
Ensemble	Boosted Trees	30.2
	Bagged Trees	85.8
	Subspace Discriminant	91.4
	Subspace KNN	79.3
	Rusboosted Trees	22.3
Neural Network	Narrows Neural Network	70.8
	Medium Neural Network	85.2
	Wide Neural Network	89
	Bilayered Neural Network	63
	Tripping Neural Network	55.7
Kernel Aproximation	Svm Kernel	74.9
	Logistics Regression Kernel	63

Since there are 100 people in total, the error matrix is very large. The error matrix is given below, but the values in the table cannot be read. Faulty points are shown with different colors. Out of a total of 3955 samples, only 339 were misclassified. This shows that a very high classification success has been achieved.

C. Gender Classification Findings

29 different algorithms were used to classify the images of the inner surface and outer surface of the hand as male or female.

Bagged, one of the highest accuracy ensemble algorithms It was obtained as 92.6% with the Trees algorithm. When the classification successes of other algorithms are examined, it is seen that many algorithms have an accuracy rate of over 90%. The results obtained in the table below are given as a percentage with the names of the algorithms. The row with the highest success rate is marked in color.

TABLE IV
GENDER CLASSIFICATION RESULTS

Algorithm Group	Algorithm Name	Accuracy rate (%)
Decision Trees	Fine Tree	79.7
	Medium Tree	73.7
	Coarse Tree	66.8
Logistics Regression	Logistics Regression	79.2
Naive Bayesian	Kernel Naive Bayesian	69.1
SVM	Linear SVM	79.8
	Quadratic SVM	90.8
	Cubic SVM	93.7
	Fine Gaussian SVM	76.5
	Medium Gaussian SVM	89.4
	Coarse Gaussian SVM	70.7
Nearest Neighbor	Fine KNN	94.7
	Medium KNN	90.9
	Coarse KNN	77.2
	Cosine KNN	91.1
	Cubic KNN	86
	Weighted KNN	92.7
Ensemble	Boosted Trees	87.5
	Baggaed Trees	92.6
	Subspace Discriminant	83.7
	Subspace KNN	92.4
	Rusboosted Trees	84.4
Neural Network	Narrows Neural Network	90
	Medium Neural Network	91.2
	Wide Neural Network	91.8
	Bilayered Neural Network	90.3
	Tripping Neural Network	88.9
Kernel Aproximation	Svm Kernel	88.1
	Logistics Regression Kernel	84

Baggaed in gender classification The error matrix of the results obtained by the Trees algorithm is given in the table below. Here, it is seen that the distribution of the data is not uniform and there are more samples in female data. However, the classification algorithm performed well and achieved a high accuracy rate. In total, only 291 samples were misclassified. Of these, 121 were misclassified as male when they were female, and 170 were misclassified as female while male.

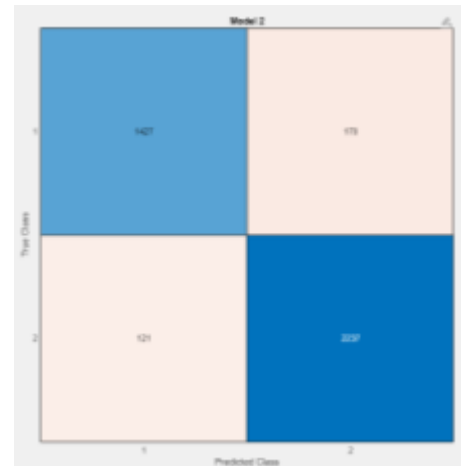


Fig. 7. error matrix of Baggaed trees algorithm for sex classification

D. Age Classification Findings

The ages of the people vary between 10 and 71 in different numbers. The distribution of these values has been given in the previous chapters. Although this distribution is not close to each other, the algorithms in the Matlab program have shown a high level of learning. Wide under the category of The Neural Network algorithm achieved a successful classification rate of 85.9%. There have also been different algorithms that produce learning results close to this value. But at the same time, algorithms that produce results with a very low success rate attract attention. The obtained results are given in the table below in a grouped way. The row with the highest success rate is marked in color.

TABLE V
AGE CLASSIFICATION RESULTS

Algorithm Group	Algorithm Name	Accuracy rate (%)
Decision Trees	Fine Tree	46.4
	Medium Tree	28.4
	Coarse Tree	16.1
Naive Bayesian	Kernel Naive Bayesian	43.5
SVM	Linear SVM	73.4
	Quadratic SVM	84
	Cubic SVM	84.7
	Fine Gaussian SVM	20.3
	Medium Gaussian SVM	72.7
	Coarse Gaussian Svm	39.8
Nearest Neighbor	Fine KNN	84.4
	Medium KNN	79.7
	Coarse KNN	52.1
	Cosine KNN	77.5
	Cubic KNN	66.9
	Weighted KNN	82.9
Ensemble	Boosted Trees	37.8
	Baggaed Trees	84.9

	Subspace Discriminant	79.3
	Subspace KNN	80.7
	Rusboosted Trees	20.8
Neural Network	Narrows Neural Network	62.6
	Medium Neural Network	79.7
	Wide Neural Network	85.9
	Bilayered Neural Network	57.9
Kernel Aproximation	Tripping Neural Network	52.7
	Svm Kernel	73.5
	Logistics Regression Kernel	51.2

It was stated in the previous titles that the age values are more between 10 and 25. For this reason, when the error matrix given in the figure below is examined, it will be seen that the erroneous estimations are concentrated here. However, despite all this, only 518 out of 3955 samples were misclassified. The remaining sample was all correctly classified.

IV. DISCUSSION AND CONCLUSION

In this study, it was aimed to determine who the people are, how old they are and what gender they are from the inner and outer surface images of their hands. For this purpose, 3955 images were collected from 100 different people. The features of these images were extracted with the median robust extended local binary model (MRELBP) method and classification was performed with many different algorithms in the Matlab program. While an accuracy rate of 91.4% was obtained for person recognition, a success rate of 92.6% was obtained for gender. Finally, a classification success rate of 85.9% was achieved for age. All these results show that it is possible to make recognition from the inner and outer surface images of the hand with high success with the method here. In the examinations made in the literature, no source directly similar to the study here was found. However, 99.7% person recognition was made with the VGG19 transfer learning algorithm on the vein images on the outer surface of the hand [33]. Although this study considers the outer surface of the hand, it is completely different from the study here. It is thought that with the effect of the transfer learning algorithm used, it may have obtained higher results than the study here. In a study conducted by Afifi [34], gender recognition study was carried out from hand images using deep learning methods. Both the inner surface and the outer surface of the hand were analyzed separately. It achieved a classification success of 94.2% in gender recognition for the inner surface of the hand and 97.3% in gender recognition for the outer surface of the hand. It is thought that this study produced higher results than the study here, because it evaluates the inner and outer surfaces of the hand separately. Al-johania and Elrefaei [35], applied deep learning algorithms on data with 500 images in one of two different datasets and 1575 images in the other, and obtained a 100% recognition rate for person recognition and a 99.25% recognition rate in another dataset. In this study, it is thought that the higher data and the feature extraction method used

cause this high recognition rate. There are many studies for person recognition from hand vein images [36-40]. However, since these studies are not directly related to the present study, they have not been compared in detail.

ACKNOWLEDGMENTS

This study was presented as a master's thesis at Kırşehir Ahi Evran University, Institute of Science and Technology.

REFERENCES

- [1] M. M. J. A. S. E. J. Fahmy, "Palmprint recognition based on Mel frequency Cepstral coefficients feature extraction," vol. 1, no. 1, pp. 39-47, 2010.
- [2] M. A. Alsmirat, F. Al-Alem, M. Al-Ayyoub, Y. Jararweh, and B. Gupta, "Impact of digital fingerprint image quality on the fingerprint recognition accuracy," *Multimedia Tools and Applications*, vol. 78, no. 3, pp. 3649-3688, 2019.
- [3] H. Chen and B. Bhanu, "Contour matching for 3D ear recognition," in *2005 Seventh IEEE Workshops on Applications of Computer Vision (WACV/MOTION'05)-Volume 1*, 2005, vol. 1: IEEE, pp. 123-128.
- [4] Ž. Emeršič, V. Štruc, and P. Peer, "Ear recognition: More than a survey," *Neurocomputing*, vol. 255, pp. 26-39, 2017.
- [5] M. Fischer, M. Rybníček, and S. Tjoa, "A novel palm vein recognition approach based on enhanced local Gabor binary patterns histogram sequence," in *2012 19th International Conference on Systems, Signals and Image Processing (IWSSIP)*, 2012: IEEE, pp. 429-432.
- [6] C. Wilson, *Vein pattern recognition: a privacy-enhancing biometric*. CRC press, 2010.
- [7] W. Shu and D. Zhang, "Automated personal identification by palmprint," *Optical Engineering*, vol. 37, pp. 2359-2362, 1998.
- [8] A. Michele, V. Colin, and D. D. J. P. C. S. Santika, "Mobilenet convolutional neural networks and support vector machines for palmprint recognition," vol. 157, pp. 110-117, 2019.
- [9] X.-Q. Wu, K.-Q. Wang, and D. Zhang, "Wavelet based palm print recognition," in *Proceedings. International Conference on Machine Learning and Cybernetics*, 2002, vol. 3: IEEE, pp. 1253-1257.
- [10] S. F. ABD RAZAK, "IMAGE ANALYSIS OF PALM (PALMISTRY)(HEALTH AND CHARACTERISTICS)," BACHELOR OF ENGINEERING, ELECTRICAL & ELECTRONICS ENGINEERING, Universiti Teknologi PETRONAS, TRONOH, PERAK, 2006.
- [11] J. Saint-Germain, *Lovers Guide to Palmistry: Finding Love in the Palm of Your Hand*. Llewellyn Worldwide, 2008.
- [12] M. M. Houck and J. A. Siegel, *Fundamentals of forensic science*. Academic Press, 2009.
- [13] A. R. Jackson and J. M. Jackson, *Forensic science*. Pearson Education, 2008.
- [14] A. Younesi and M. C. J. P. C. S. Amirani, "Gabor filter and texture based features for palmprint recognition," vol. 108, pp. 2488-2495, 2017.
- [15] A. K. Jain, A. Ross, and S. Prabhakar, "An introduction to biometric recognition," *IEEE Transactions on circuits and systems for video technology*, vol. 14, no. 1, pp. 4-20, 2004.
- [16] Y. Bulatov, S. Jambawalikar, P. Kumar, and S. Sethia, "Hand recognition using geometric classifiers," in *International Conference on Biometric Authentication*, 2004: Springer, pp. 753-759.
- [17] L. Fei, B. Zhang, W. Zhang, and S. J. I. S. Teng, "Local apparent and latent direction extraction for palmprint recognition," vol. 473, pp. 59-72, 2019.
- [18] J. Cui, "Multispectral fusion for palmprint recognition," *Optik-International Journal for Light*, vol. 124, no. 17, pp. 3067-3071, 2013.
- [19] S. Zhang and X. Gu, "Palmprint recognition method based on score level fusion," *Optik-International Journal for Light and Electron Optics*, vol. 124, no. 18, pp. 3340-3344, 2013.
- [20] S. Zhang and X. Gu, "Palmprint recognition based on the representation in the feature space," vol. 124, no. 22, pp. 5434-5439, 2013.
- [21] J. Li, J. Cao, and K. Lu, "Improve the two-phase test samples representation method for palmprint recognition," *Optik*, vol. 124, no. 24, pp. 6651-6656, 2013.
- [22] S. Zhao and B. Zhang, "Deep discriminative representation for generic palmprint recognition," *Pattern Recognition*, vol. 98, p. 107071, 2020.

- [23] W. M. Matkowski, T. Chai, and A. W. K. Kong, "Palmprint recognition in uncontrolled and uncooperative environment," *IEEE Transactions on Information Forensics Security*, vol. 15, pp. 1601-1615, 2019.
- [24] Z. Liu, Y. Yin, H. Wang, S. Song, and Q. Li, "Finger vein recognition with manifold learning," *Journal of Network Computer Applications*, vol. 33, no. 3, pp. 275-282, 2010.
- [25] D. Hong, W. Liu, J. Su, Z. Pan, and G. Wang, "A novel hierarchical approach for multispectral palmprint recognition," *Neurocomputing*, vol. 151, pp. 511-521, 2015.
- [26] M. D. Bounneche, L. Boubchir, A. Bouridane, B. Nekhou, and A. Ali-Chérif, "Multi-spectral palmprint recognition based on oriented multiscale log-Gabor filters," *Neurocomputing*, vol. 205, pp. 274-286, 2016.
- [27] F. Ma, X. Zhu, C. Wang, H. Liu, and X.-Y. Jing, "Multi-orientation and multi-scale features discriminant learning for palmprint recognition," *Neurocomputing*, vol. 348, pp. 169-178, 2019.
- [28] D. Tamrakar, P. J. J. o. V. C. Khanna, and I. Representation, "Kernel discriminant analysis of Block-wise Gaussian Derivative Phase Pattern Histogram for palmprint recognition," vol. 40, pp. 432-448, 2016.
- [29] D. Hong, W. Liu, X. Wu, Z. Pan, and J. Su, "Robust palmprint recognition based on the fast variation Vese-Osher model," *Neurocomputing*, vol. 174, pp. 999-1012, 2016.
- [30] C. Turan and K.-M. Lam, "Histogram-based local descriptors for facial expression recognition (FER): A comprehensive study," *Journal of visual communication and image representation*, vol. 55, pp. 331-341, 2018.
- [31] T. TUNCER and A. J. T. B. V. B. v. M. D. Engin, "Yerel İkili Örüntü Tabanlı Veri Gizleme Algoritması: LBP-LSB," vol. 10, no. 1, pp. 48-53, 2017.
- [32] L. Liu, S. Lao, P. W. Fieguth, Y. Guo, X. Wang, and M. J. I. T. o. I. P. Pietikäinen, "Median robust extended local binary pattern for texture classification," vol. 25, no. 3, pp. 1368-1381, 2016.
- [33] H. Wan, L. Chen, H. Song, and J. Yang, "Dorsal hand vein recognition based on convolutional neural networks," in *2017 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, 2017: IEEE, pp. 1215-1221.
- [34] M. Afifi, "11K Hands: Gender recognition and biometric identification using a large dataset of hand images," *Multimedia Tools and Applications*, vol. 78, no. 15, pp. 20835-20854, 2019.
- [35] N. A. Al-johania and L. A. Elrefaei, "Dorsal hand vein recognition by convolutional neural networks: Feature learning and transfer learning approaches," *International Journal of Intelligent Engineering and Systems*, vol. 12, no. 3, pp. 178-91, 2019.
- [36] S.-J. Chuang, "Vein recognition based on minutiae features in the dorsal venous network of the hand," *Signal, Image and Video Processing*, vol. 12, no. 3, pp. 573-581, 2018.
- [37] S. A. Radzi, M. K. Hani, and R. Bakhteri, "Finger-vein biometric identification using convolutional neural network," *Turkish Journal of Electrical Engineering & Computer Sciences*, vol. 24, no. 3, pp. 1863-1878, 2016.
- [38] L. Wang, G. Leedham, and S.-Y. Cho, "Infrared imaging of hand vein patterns for biometric purposes," *IET computer vision*, vol. 1, no. 3-4, pp. 113-122, 2007.
- [39] L. Wang, G. Leedham, and D. S.-Y. Cho, "Minutiae feature analysis for infrared hand vein pattern biometrics," *Pattern recognition*, vol. 41, no. 3, pp. 920-929, 2008.
- [40] O. O. Iroanya, T. F. Egwuatu, O. T. Talabi, and İ. S. Ogunleye, "Sex Prediction Using Finger, Hand and Foot Measurements for Forensic Identification in a Nigerian Population". *Sakarya University Journal of Science*, vol. 24 no. 3, pp. 432-445, 2020.

BIOGRAPHIES



EMRAH AYDEMİR Elazığ, in 1987. He received the M.S. degrees in computer teaching from the University of Elazığ Firat, in 2012 and the Ph.D. degree in informatics from Istanbul University, Turkey, TR, in 2017. From 2012 to 2015, he was an Expert with the Istanbul Commerce University. Since 2017, he has been an Associate Professor with the management information systems, Sakarya University. He is the author of three books, more than 10 articles, and more than 40 conference presentation. His research interests include artificial intelligence, microcontroller, database and software.

RAGHAD TOHMAS ESFANDIYAR ALALAWI She received the B.S. in computer engineering from the University in Iraq. She completed her master's degree in Advanced Technologies at Kırşehir Ahi Evran University. She works on artificial intelligence and image processing.

Smart Steps Towards Sustainable Transportation: Profitability of Electric Road System

Hasan Huseyin Coban

Abstract—The global average temperature has increased by about 1°C compared to pre-industrial times and the temperature increase continues. The effects of climate change are already affecting living conditions on Earth. Through the Paris Agreement, the countries of the world have committed to limit global warming to below 2°C compared to pre-industrial times and to make efforts to keep the increase below 1.5°C. The transportation sector accounts for about 15% of Turkey's greenhouse gas emissions, and more than three-quarters come from diesel fuel. Already in ten years, transport emissions must be reduced by 70% according to the national targets. It is about both reducing emissions by 2030 but also ensuring that we are completely emission-free by 2050. The vast majority of emissions in the transport sector come from road traffic. Electric vehicles have emerged as the most promising solution to energy and environmental security concerns. However, a major constraint for electric vehicles has been the limited range and fast charging capacity. Various electric road technologies have great potential to reduce dependence on fossil fuels, reduce greenhouse gas emissions, reduce air pollution and reduce noise in urban areas. This article proposes to discuss the inductive and conductive charging system and required system configuration in the concept of electric roads for electric buses. The model solution algorithm is designed and it is validated on the E-bus route R40 in Kayseri city, Turkey. It is found that if an electric road system is used instead of a diesel bus for a bus line route of 35.5 km, it will amortize itself in 10 years.

Index Terms—Electric Road System, Design process, Sustainability, Electric bus.

I. INTRODUCTION

THE HEAVY traffic is now increasing rapidly, contributing to ever higher greenhouse gas emissions [1]. While some people defend the transfer of people and heavy loads to rail, there are other conceivable alternatives. One is to electrify road transport instead. An electrified road system, "electric roads", or "Eroad", or "electric road system" (ERS) where electrical

energy is transferred during movement from the road to vehicles for both propulsion and charging, has great potential for reduced dependence on fossil fuels and increased energy efficiency in the transport sector.


Electric vehicles (EVs) are seen by many as a possible solution to reduce harmful emissions to the environment, reduce dependency on fossil fuels and increase energy efficiency in the transportation sector [2]. Unfortunately, most environmentally friendly energy storage systems, such as batteries, contain a lower energy density compared to fossil fuel, which has a significant impact on the vehicle's range [3], [4]. A battery large enough for long-distance transport often results in a significant increase in cost and weight, which means a reduction in possible transport volume. An alternative to battery-powered operation is to transfer energy from the road to the vehicle for both propulsion and charging during movement. Expanding the ERS between larger urban areas means that the majority of the route can be powered by electricity from the road network, and the remaining route can be powered by a smaller battery.

Road transport uses electricity as fuel and the technology used to transmit electricity also includes a properly designed infrastructure that is lacking today. This study shows that the expansion of the infrastructure for the supply of electricity to buses is not as complex and costly as the expansion of the railway.

Road freight accounts for nearly one-fifth of annual global oil consumption [5]. Today, the climate problem is one of the most current issues in society. The EU aims to reduce its emissions by 80-90% by 2050 compared to 1990 [6]. The ministry of environment and climate change aims for Turkey's net emissions to be zero by 2050 and for transportation to reduce its emissions by 70% by 2030. In addition, Turkey should also be the region's first fossil-free welfare state. To meet these demands, more and more people are switching from fossil fuels to alternative fuels, including electric cars.

The EV is nothing new, and one of the reasons it hasn't been adopted is because of the time-limited range when refueling on the go. As the vehicles are electric, there is a need for charging on long journeys. Today's market-leading passenger cars have a range of around 400km and can be quickly charged to 80% in 30 minutes at selected locations [7]. For longer, uninterrupted transport, fast charging is not an option, especially when it

HASAN HUSEYIN COBAN is with the Department of Electrical Engineering University of Ardahan University, Ardahan, 75002, Turkey, (e-mail: huseyincoban@ardahan.edu.tr).

 <https://orcid.org/0000-0002-5284-0568>

Manuscript received Jan 17, 2022; accepted Jan 28, 2023.

DOI: [10.17694/bajece.1058841](https://doi.org/10.17694/bajece.1058841)

comes to commercial traffic, as there will be too many and too long charging breaks.

The transition to battery-powered buses for electric operation adapted to both electrified roads and stationary charging represents a technological leap that could be significant in reducing carbon dioxide (CO₂) emissions from road transport. The technology is at an early stage, making profitability assessments of infrastructure investments more uncertain than they usually are. A discussion of ERS and their potential to reduce dependence on fossil fuels can be found in the report [8], [9], which is the source of this article and the investigation into fossil-free vehicular traffic.

This study aims to investigate the electrification of municipal buses in Kayseri province. Through this study, it is to lay the foundation for a project on the electrification of bus traffic in the urban area of Kayseri by highlighting the key questions to be resolved. This is expected to contribute to better air quality in the city and Turkey's achievement of its environmental targets by 2050 [10]. The other aim of this study is to investigate what technical changes should be made in the infrastructure of the city in order for electric buses (EBs) to be charged and to develop a concept for a new road design that supports the future development of electrified road technology for a bus line with intensive human transport consistent with its sustainable intentions. In addition, the economic feasibility of constructing an electric road along the bus line using small batteries will be demonstrated. This study does not cover the electrification of bus lines outside the urban area, such as rural and regional buses or school transport. In addition, of course, the cost of the actual investment in the ERS and the proportion of fuel-powered vehicles that switch to electric operation play a large role in the profitability assessment. This is described in the analyzes by varying the investment cost in the various calculations and by varying the number of fuel-powered vehicles that switch to electric operation. Electric roads are a relatively new field, and articles and reports in the literature have been used to find information [8], [11]–[14]. Relevant literature was used to find information about EBs and especially the characteristics of batteries.

Chapter 1 is a short introduction about why the study is being carried out and what is to be done. Chapter 2 is a description of the present and future bus systems in Kayseri. Routes, route length, departure frequency, and other important information are given for the selected bus line. A comparison of fuel consumption and emission rates is presented to motivate the conversion from internal combustion buses to EBs. Chapter 3 specifies all the components required to build an ERS and is discussed from an economic point of view. Chapter 6 presents all the results from the study and the possible future study topics.

II. THE PRINCIPLE AND TECHNOLOGY OF EB AND ERS

This section presents the available technology for EBs and the infrastructure needed to operate different types of EBs. Then it briefly explains the theory behind how energy can be

extracted from the road and how each technology is used to deliver energy to buses.

A. Electric Buses

In this section, focusing on city buses, diesel buses, and battery-powered buses are examined. By comparing them, energy consumption, energy storage, usage, energy transfer, and environmental effects are examined. EBs are generally divided into the following categories: electric hybrids, plug-in hybrids, and EBs. This study mainly targets EBs. The advantages and disadvantages of EBs are listed below [15], [16].

Some advantages of EBs are:

- More enjoyable trips. The quieter and smoother operation of EBs provides a more enjoyable environment for both bus drivers and passengers.

- No exhaust pipes. This is how you can limit or completely eliminate greenhouse gas emissions from the bus itself.

Some disadvantages of EBs are:

- Limited range. The range will be longer if the bus' battery is quickly charged along the route. However, using the ERS by using a smaller battery instead of a large battery will eliminate the range problem.

- In cold climates, it affects the range more directly because the batteries have lower capacity than in hot regions, and indirectly because the heating in buses requires more energy.

- Less flexibility. When using an EB, it is necessary to plan the route in advance, each charging strategy changes the range capacity of the bus. The EB, which is suitable for one bus line, may not be suitable for another line at all.

- Infrastructure is required. Charging stations are needed, they are all charged at the last stops or depot, but some have to be recharged along the way, usually with the help of pantographs at the main stops.

- Risk. The electric technology applied to buses is still relatively immature, unlike conventional Internal Combustion Engine (ICE) technology which is mature and well-understood.

- Cost of capital. Especially if the infrastructure does not exist, electrification will be very expensive. EBs are more expensive than diesel buses in terms of price, maintenance, and repair.

B. The design of the EB charging infrastructure

This question is described as one of the most important aspects when purchasing EBs. There are four main options for how charging infrastructure can be designed: depot charging, additional charging, and ERSs [17]. Charging can be done either conductively (via contact) or inductively (wireless).

Plug-in charging [18]: The battery is charged by manually connecting a cable from a charging station to the bus as a plug-in and this technology is mainly used for depot charging. The driver will need to get out of the vehicle and plug the cable into the bus, but the installation is simple and the plug is less affected by winter weather compared to other charging methods.

Pantograph charging [17]: There are different charging methods in this category. There is a mechanical arm that connects the bus to the electrical grid and this arm is controlled by the bus driver from inside the bus. Pantographs can either be

overhead or placed under the bus. If the pantograph is on the bus, it has a pair of rails to which it is attached, and the mechanical arm lifts it towards the mast. There is also the possibility that the pantograph can be placed at the bottom of the bus and connected to the rails on the ground.

Wire charging [19]: Another automatic conductor charging method is to charge the bus via a wired system suspended above the road. It is a concept similar to a trolleybus or tram, except those buses can automatically connect and disconnect from this system and run freely on battery power for a while.

Inductive charging [18]: The last category is wireless inductive. Here charging takes place wirelessly along the way, either while standing still or while working where the induction comes from coils etched about 5-10cm below the road. This can have large transmission losses in the charging concept and is an immature technology concept.

The amount of time needed to successfully connect the desired electrical capacity to the facility depends on many different parameters. Care should be taken to ensure that detailed plans, terrain, transformers, etc. are ready and suitable for the purpose. Depot charging and plug-in are the most common in the world today and the system that is the cheapest to introduce and requires the least adaptation of the infrastructure in the urban area. It is also more flexible as the buses can run on all routes. The main disadvantage is that it requires larger batteries with a higher capacity and it is not certain that the battery will last a full day of driving.

C. Typical charge effects

EBs adapted to parked or end-stop charging require very large batteries to travel as far as possible on a single charge; which usually means a more expensive purchase price for these buses, but cheaper and simpler infrastructure as all charging takes place.

TABLE I
INSTALLATION PRINCIPLES AND CHARGING EFFECTS FOR
DIFFERENT CHARGING TECHNOLOGIES

Charging technology	Power (kW)	Installation
Electric road system	Conductive: 50 - 60	Electrical panels are needed at some connection points along the bus line.
	Rail in road: 200 - 800	
	Inductive: 25 - 100	
Plug-in	30 - 150	Common grid station for the depot. Connection with cable at each depot location.
Pantograph	300 - 700	Charging at a post, beam, or gantry either at the end stop, terminal, or dedicated charging point in the vicinity of the end stop.
Induction (Stationary)	200	Charging plate under the ground (asphalt).

Charging in the depot can be done in several different ways, but plug-in charging is generally used because buses charge

longer at night. In case any bus needs additional charging during the day, it may also be necessary to have a few quick chargers in the tank. Longer distances and high-frequency city traffic that require charging during the day can be difficult to implement tank charging. This is mainly because it means more empty driving during the day as buses need to be recharged. It can also mean extra buses that replace buses that need to be charged.

Table 1 shows the typical power ranges of different charging technologies and their applications around the world. The effects in Table 1 can be adjusted for how long the buses need to be charged.

D. Review of Electric Road System

Ten years ago, when ERS systems [8], [11] began to be discussed, the world was focused on developing batteries. Some investors saw it a little differently, they wanted a solution that could connect vehicles to the network 24/7 and saw the road as a solution. Charging is a matter of time, so the most natural way is to turn the road into a charging asset. There are 1.4 billion vehicles on the planet, and there is no way to provide one charger per vehicle and this is not sustainable [20]. It needs a common charging platform and the key is inductive or conductive charging via ERS. Electric road technology can charge any electric vehicle, bus, truck, or taxi while driving or parking in any weather conditions. Because vehicles do not need dedicated downtime for charging, the fleet needs to address various industry challenges, such as increasing operational uptime. It also increases the load range, which can be basically unlimited, balancing the demand in the electricity grid by providing day and night charging. The participation of several different actors and systems is needed for electric roads to become a functioning transport solution that produces good climate effects (see Figure 1).

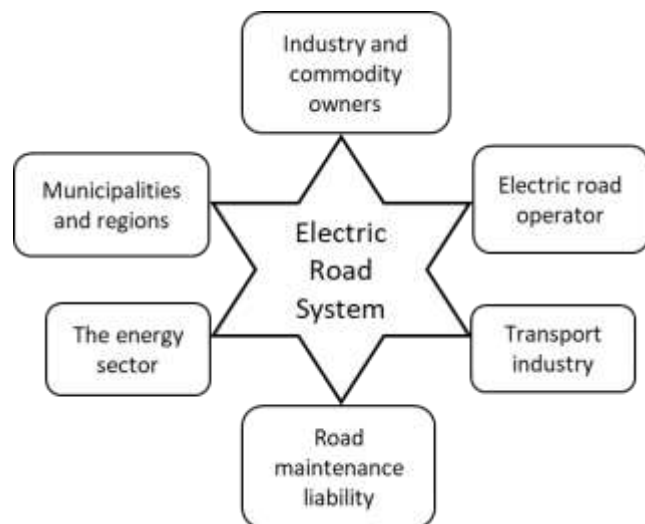


Fig. 1. The system of ERS

There are different technologies for electric roads today that can be divided into 3 main types, based on energy transfer from a road to a vehicle [8], [19]:

- inductive transmission via electromagnetic fields from the road body,
- conductive transmission via overhead line,
- conductive transmission via tracks or conductors in the road.

Induction is a way of wirelessly transferring energy between the source and receiver. The electric road uses induction to charge EVs moving on it, via receivers mounted on the underside of the vehicles. A system common in our households; wireless phone chargers and induction cookers are examples. Induction technology is used to charge the electric vehicle in the same way as the phone, it is done by static induction, which means that both the source and receiver are stationary during the transfer of energy. However, what is under development is to use similar technology with the difference that the receiver (vehicle) is in motion and thus dynamic induction is achieved. This technology works considerably the same as static induction charging, the challenge lies in ensuring safety.

Conduction is the opposite of induction and deals with the transfer of energy through physical contact between the source and the receiver. There are three ways to charge vehicles conductively while driving; power supply above the road, power supply beside the road, and power supply on the road. This means that energy is transferred from a charging rail in the road to the vehicle via a folding collector under the vehicle. The principle can be compared to the electric propulsion of trains, a lever is opened and the vehicle is in contact with the power source at the same time as it is in motion. One challenge with the transmission - especially in pantographs on the road - is that trucks and cars have large height differences. A passenger car will need a long arm to reach the same powerline from which the trucks will be powered.

There are big differences between EBs adapted for different charging strategies. The different available ERSs for EBs are shown in Figure 2 and can be divided into 3 different categories. These categories are used by different countries around the world (see Table 3) to help cities with different conditions charge EBs.

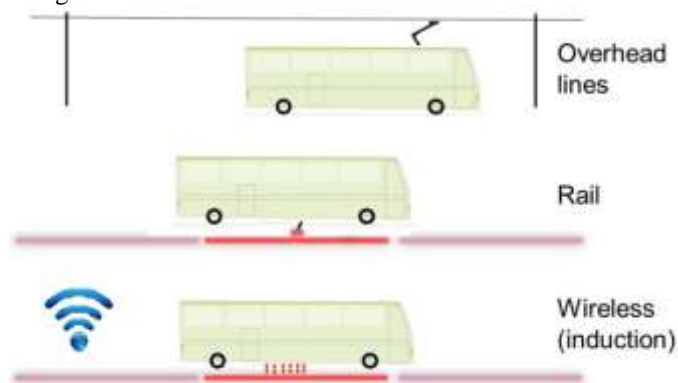


Fig.2. The three main concepts for road ERS and energy transfer to moving EBs.

In the ERS, instead of charging the bus by stopping, they can be charged while driving along the road, and instead of a large battery, a battery of about 70 km is used. If the bus is connected to cables suspended above the road, the bus can be

automatically disconnected from this cable if necessary, during operation. The battery is activated when it leaves the electric road and then easily reconnects to the line when it returns to the ERS. The bus can be charged via induction, meaning you can have small batteries on buses and not have to stand still to recharge. Induction is a very expensive alternative to build in and maintain. It costs around 1-2.2 million \$ to pave an induction road in a rural setting.

A vehicle that uses a future network of electric roads will have a long-range, reduced need for battery which means low cost, and reduced need to stop to charge which means flexibility. The road to electrification will certainly lead to a better and greener future.

The electrified roads of the future will need some form of payment system to charge for infrastructure and energy use. How such payment systems should be designed is currently still undecided and has been explored in some limited studies [8], [13] prior to this study. Payment systems need to be able to handle complex use cases with multiple actors, roles, and commercial relationships. EVs on electrified roads can be treated similarly to road tolls today, but with a supplement for electricity consumption. For a study of toll systems for electrified roads, lessons should also be drawn from existing toll systems for rail traffic. The payment system to be integrated into the vehicle while driving continuously reports location data and time, including vehicle information, and sends this information via mobile network to be processed for tax calculation and payment according to tariffs and price lists. Road tax is a way to finance future maintenance and investment in road infrastructure essential for economic growth. Numerous countries in Europe have implemented different types of systems on roads to generate tax revenue. The basis for collecting road tax is very similar to the information a payment system would need for electrified roads.

E. The investment cost of the ERS

Uncertainty about what electrified roads could cost is significant. In Sweden, some companies assess that in the long run, road-rail feeding could result in an investment cost of approximately \$1 million per km. However, this may come across as an exaggerated observation. In Germany, quite high costs are observed for the conductor feeding alternative to the highway. In this study, it has started with the option of conductor feeding from the mains to the highway, which is the most advanced technology today. In the applications and studies, it is stated that the cost of ERS powered from the grid is between 1 and 3 million \$ per km in both directions. Therefore, in this study, it is assumed that the cost of two-way roads is \$1.6 million per km (see Table 3). A relatively short economic life is assumed for the electrified road, only 15 years. Therefore, the annual maintenance cost is assumed to be small, except for the wear and tear cost incurred when using an electrified road. Any additional costs for the expansion of electrical grids required to ensure that the required electricity reaches the road are beyond the scope of this work. In addition, even if the operation of heavy traffic can be solved with batteries, the capacity in the network will need to be expanded along the highways. In this study, it is assumed that buses

should invest in a battery that provides a range of 70-100 km. Table 2 presents the costs of implementing ERS technology for each solution.

F. Earlier studies and projects

Today, there are a limited number of companies that can build in ERSs. The technology developed by the Elways [21] company is based on ground-based conductive charging. The basic cost of investing in an electric road with Elway's technology is estimated at approximately \$ 0.5 million per kilometer for both driving directions. Sweden-based company Elonroad's [22] technology relies on a track on the road that feeds batteries in EVs passing by, provided they have a receiver that sticks to the bottom of the vehicle.

An electric road is supplied with electricity from a surrounding power grid, and there is switchgear along the path that forms the boundary between low voltage and medium voltage.

TABLE II
ERS COST COMPARISONS

Type of ERS	Cost (M\$ per km)	Maintenance Cost (\$per km)
Inductive	2,2	33.000
Conductive Overhead	1,4	16.000
Conductive Rail	0,8	10.000

TABLE III
OVERVIEW OF COMPLETED AND ONGOING VARIOUS ELECTRIFIED ROAD PROJECTS

Name	Reference	Country	Type of ERS	Company	Comments
Tel Aviv	[27]	Israel	Inductive	ElectReon and Dan Bus Company	600 meters of electrified road along a two-kilometer route between Tel Aviv University Railway Station and Klatzkin Terminal in Ramat Aviv
Smartroad Gotland	[5]	Sweden	Inductive	NA	It installed 1.6 km of electric road on a 4.1 km section between the city of Visby and Visby Airport.
EVolution Road	[28]	Sweden	Inductive	Elonroad, University of Lund	One-kilometer-long section of the road will be equipped with electric rails, The total investment is 9M€. It charges a city bus while driving with 150kW
VICTORIA	[29]	Spain	Inductive	Circe	Eight, 80 cm long, 50 kW coils are installed along the 100-meter-long road, in Malaga city
Select	[30]	USA	Inductive	Siemens	The work was done under the direction of the South Coast Air Quality Management District. Consortium with more than 10 universities.
ELinGO	[31]	Norway	Conductive	NA	In 2018, a preliminary study on the electrification of the E39 was carried out
Visby	[32], [33]		Inductive	ElectReon	the first fully operational electric bus utilizing the company's wireless charging infrastructure.
Michigan	[34]	USA	Inductive	Electreon	ERS to pilot on a 1-mile section of road in Detroit first in the United States.
Wave	[35]	USA	Inductive	Wave	A system in the range from 25 kW to 40 kW can be tested using a 20-seat passenger bus
Karlsruhe	[36]	Germany	Inductive	ElectReon	a bus line connecting the new EnBW training center in Karlsruhe's Rhine harbour to the local public transport system hub.
Brebemi	[37]	Italy	Inductive	ElectReon	The Brebemi highway in Lombardy, 1km. Iveco Bus E-Way and Fiat 500e tested recharge while driving at around 70 km/h.
FABRIC	[38], [39]	Italy	Inductive	ICCS, ERTICO, TRL, ICCS, KTH	outside Torino city, a 260m test track developed. 25cm, 20-100kW
FABRIC	[39]	France	Inductive	ICCS, ERTICO, TRL, ICCS, KTH	The project will allow relevant experiments associated with dynamic inductive charging use on road in Satory.
e-Road Arlanda	[40]	Sweden	Conductive	Elways	It is completed 1.25 miles of the ERS on a public highway, with plans to expand to other parts of the country.
Milton Keynes	[41]	UK	Inductive	Mitsui and Arup	Line-7 15-mile route, covering two suburbs of Wolverton and Bletchley.
Inter-city highway	[41]	Germany	Inductive	ElectReon	Near Cologne city, the investment is €1.9 million, 100m of ERS
E-highway	[42]	Germany	Conductive	Siemens, SPL Powerlines	The system on the Bundesstraße 462 (B 462) federal highway, a 3.4-kilometer route
OLEV	[43]	South Korea	Inductive	OLEV	a bus line in the city of Gumi, Sejong, and Gumi

Currently, there are many completed or ongoing ERS development projects (OH-Lkw [23], COILERS [24], [25], StratON [26]) around the world. The electric road is a promotional project in Sweden to test a new sustainable alternative for heavy traffic with the potential to replace fossil fuel. In 2016, the world's first electrified road between Sandviken and Kungsgården was inaugurated. The ERS used is conductive electricity transmission over overhead power lines, this technology has long been used and proven for trams. The Swedish Transport Administration is carrying out important studies for electrified roads across the country. Inductive energy transfer is available for part of a bus line in Gothenburg and Södertälje. Consortium members are AB Volvo, Bombardier, Vattenfall, Södertälje municipality, Telge Energi, and Göteborg Energi. For trucks, there is an overhead line built in Gästrikland along a road of approximately 2 km between Storvik and Sandviken, and approximately 1 km of electrified roads between Arlanda and the logistics center in Rosersberg. Consortium members are WSP Analysis and Strategy, Kilenkryset, Elways, NCC, e-Traction, Arlandastad Holding, Airport City Stockholm, Royal Institute of Technology in Stockholm KTH, Siemens, Scania, Ernst Åkeri. German highways usually have one lane where only trucks travel. For this reason, the overhead line was chosen for the project that did not have a car and truck combination. Overhead power transmission by Siemens in Berlin and California, testing of OLEV's inductive power transmission solution in South Korea and work in the UK are ongoing. Current technical conditions and developments for inductive and conductive ERS technologies are summarized in Table 3.

III. RESEARCH CONTEXT AND LIMITATIONS

This chapter provides a general introduction to how the ERS system for EBs works and which parameters are important to it. It offers the existing technology for EBs and the infrastructure needed to run different types of EBs. Finally, it is followed by a subsection that presents the assumptions used in this study. The assumptions are presented in a numbered list with brief explanations of why they are used.

We like to avoid fossil-fueled energy sources, and rail transport is on the rise. There is approximately 32509 km (42%) of highways and main roads in Turkey (as of December 2022). The transportation sector accounts for about 15% of Turkey's greenhouse gas emissions, and more than three-quarters come from diesel fuel [44]. Some municipalities in Turkey have already started using EBs to help them meet their environmental goals by 2050. Transportation with EBs in Istanbul [45], and Samsun [46] started in 2014, and 2022 respectively. Air quality in Kayseri province is considered one of the biggest sources of CO₂ equivalent gases emitted by heavy vehicles.

Turkey's vehicle fleet caused around 81 million tons of greenhouse gas emissions in 2020, with about a fifth of emissions coming from road transport [44]. One of the environmental goals of the Ministry of Environment and Climate Change is that greenhouse gas emissions from transport should be reduced by at least 70% compared to

2010, by 2030 at the latest. Turkey also has a shared responsibility with other countries for the global goal of reducing greenhouse gas emissions to zero in the long term, which must be pursued and achieved. How this will be achieved is not yet known, but priority is given to having a fossil fuel-neutral vehicle fleet by 2030. Existing policy tools and measures are not enough to achieve the goal, more must be done. From a sustainability standpoint, it is absolutely critical that the transition to EVs takes place on time while maintaining flexibility for future systems and that transaction costs are as low as possible. The development of electric transport systems is clear, but how energy is provided or transported in the vehicle is still unclear. Although there are various solutions and variations on this subject with different applications, the ability to bring technology development solutions to the market will provide the main benefit. Although many governments around the world, especially the European ones, have developed projects, no investment has been made in any projects about the electric road in Turkey yet.

A. Energy requirements for EBs

Estimated energy consumption for EBs varies between available articles [34], [47], [48]. Energy use depends on many different parameters such as passenger load, topography, driver behavior, road quality, and climate. Therefore, the energy consumption in Kayseri will vary according to the season. In this study, it is assumed that the municipal buses used in Kayseri city are 12 meters long. Currently, the fuel type and the number of used buses is CNG (Natural Gas) and diesel, 253 and 360, respectively [49]. A reasonable assumption is that such a bus draws around 1.5 kWh/km. If the bus is equipped with a heat pump for cold seasons, it is assumed that about 0.5 kWh/km is used for the heating system and used for days with temperatures as low as 7°C. It is assumed that the locations of the stops are located at a distance of approximately 350 m from each other.

B. Ranges and battery sizes

Since different charging strategies have different requirements for how far the buses have to travel, different battery sizes are provided depending on the charging strategy applied to EBs. It is assumed that the depot-charged buses have large batteries (battery capacity of about 420 kWh). Batteries should never be completely discharged. State of charge (SOC) is an expression of how much of the energy stored in the battery can be reused before the battery needs to be charged. In Table 4, it is assumed that the SOC value of battery-powered buses is 70%.

The ambient temperature in Kayseri varies between -32°C and +40°C, under cold climate conditions, it directly affects the capacity of the battery and this causes the vehicle's range to be shortened. For this reason, it is more advantageous to use a small battery in the EB and supply energy from the ERS.

TABLE IV
ASSUMPTIONS FOR CALCULATIONS

Parameter	Assumed Value
Energy consumption standard 12 meters bus	1.5 kWh/km
	also approx. 0.5 kWh/km for the heating system
SOC (State of Charge) range - end-stop charged buses	50%
SOC (State of Charge) range - depot-charged buses	75%
Battery Size - Terminal charged buses	200 km
Battery size - Depot-charged buses	400 km
Battery size - Plug-in charged buses	200 km
Battery size - ERS	70 km

C. Annual costs and Cost calculation

The purchase price of the EB is higher than the conventional bus, in this study it is assumed that the price of a 12m EB and a fossil fuel bus is approximately \$362.000 and \$181.000, respectively. If the battery price is not taken into account, the difference in the purchase price between the Pantograph-charged bus and the depot-charged bus is negligible.

The annual cost of an EB is higher than a diesel bus, but from a socio-economic point of view when noise and air pollution are valued, EBs are profitable. In addition, EBs have significantly lower energy costs due to the more energy-efficient use of electric motors and the fuel price. The current fuel price per kilometer cost for EBs is assumed to be \$0.18 (2 kWh/km). The costs for a community bus are summarized in Table 5.

Battery EBs can lead to high driver costs as they have a limited range and possibly cannot operate continuously throughout the 24-hour service and will have to return to the depot to change buses at some point in the day. An assumption of 0.25\$/km and 388.725 kilometers traveled gives an assumption of \$97.182 total. Important items of the total cost are vehicle purchases, battery, and maintenance costs. Battery sizes can be optimized to reduce costs. Unfortunately, it is not clear how long the infrastructure can be used, as charging technology is currently relatively young.

D. Heating system in the bus

An ICE-powered bus heats the cabin with excess heat from the engine, but an EB that needs cabin heating will use most of the electrical energy pushing it forward. Air source heat pumps can be used to regulate the energy use of the bus air conditioning system up to outside temperatures of around 7°C or less.

E. The bus depot in Kayseri

At the Kayseri bus depot, many of the buses are parked outdoors, without a roof. EVs have a battery thermal management system to keep the battery pack between optimum operating temperatures in hot and cold climates.

However, when the vehicle is parked, extreme weather conditions can also affect the battery life. Years are not the only thing that ages a battery. While time is the main cause of deterioration for electric vehicle batteries, the state of charge of the battery and exposure to extreme weather conditions also significantly affect battery life. Figure 3 shows the parking area of the bus depot from Google Maps.

TABLE V
ASSUMPTIONS FOR COST CALCULATION

EB Consumption	EB type	12 meters	0,8-1,6 kWh/km
		18 meters	1,9-3,0 kWh/km
Bus Cost (ICE)	181.000 \$		
Electric bus cost	362.000 \$		
Diesel Consumption	30 liter /km (in the city)		
Maintenance Cost	0.3 \$/km		
Battery size	Big battery: 300 kWh		
	Small battery: 120kWh		
Range	Big battery: 200 km		
	Small battery: 70km		
Maintain charger	3% of investment (\$/year)		
Cost battery	150 \$/kWh		
Fuel cost	0.09 \$/kWh (1.6TL/kWh)		
Fuel cost	1.22 \$/liter (23TL/liter)		
Number of runs	388.725 km/year		
Battery lifetime	5 years		
ERS lifetime	15 years		



Fig.3. A bird's-eye view of the existing bus depot in Kayseri.

F. The terrain of Kayseri city

The terrain of Kayseri is hilly in some places. Too much uphill driving will shorten the range of the buses. Fluctuations in topology can have a huge impact on the range performance of EBs. The city is not very hilly although there are some elevations, the topography was studied using Google Earth. Figure 4 shows the city bus routes' from Anafartalar to Ebic Dadagi in Kayseri on a map.

G. Diesel Fuel Price

In order to predict the price of diesel in 2030, we need to make assumptions about how the price of fossil diesel will develop by then. In the first quarter of 2023, the taxed diesel price was \$1.2 (23TL) per liter. Global warming, the coronavirus epidemic, and the Russia-Ukraine war have led to an increase in fuel prices. The industry remains uncertain about the future market for diesel fuels, as some European states have passed a law banning the sale of new petrol and diesel cars in 2030. A comprehensive study is needed to achieve the ban on diesel vehicles, it is not yet clear whether the investments will be realized. Therefore, it is assumed that in 2030 the diesel price will be \$2 per liter including all taxes.

IV. METHODOLOGY AND DATA GATHERING

This section describes the research process, research design, and what methods were used to collect data. It also describes reliability and validity, and how these were taken into account in the study process.

A. Description of Possible Pilot Route

Line R40 runs in both directions, so the graph in Figure 4 should read left to right when looking at the route from Anafartalar to Ebic Dadagi. The route is 35.5 km long and the average value of the land slope is %4. The route has 73 stops and shown in Figure 4 shows that it will be implemented by the ERS infrastructure. Figure 5 shows the topography of the R40 bus line.

B. Choice of Technology

For electric cars, the unit kWh is often used to describe the capacity of the electric car battery, that is, how much electrical energy can be stored, and a capacity of about 40 kWh is common. With a 40-kWh battery of which 90% is available, and with a consumption of 0.17 kWh/km, the range is roughly 210 km. In EVs, the battery is the most expensive component, making them generally more expensive than fuel cars. Electric cars are expected to get cheaper as the cost of batteries for electric cars falls [50]. The price of lithium-ion batteries for EVs dropped from around \$1000 per kWh in 2010 to just over \$200 per kWh for battery packs in 2016 and continues to fall. By the year 2030, the price is estimated to be able to drop to \$100 per kWh [50]. The distribution of electrified road distances, battery size, and the lifetime of the batteries are closely linked. In the absence of an electrified road, large batteries are needed, partly to extend the range and at the same time preserve the life of the batteries. The more electric distances it travels, the more the vehicle can be charged during the journey, meaning

the batteries do not have to be the same size. At the same time, the life of the batteries can be increased by a higher percentage of electric trips.

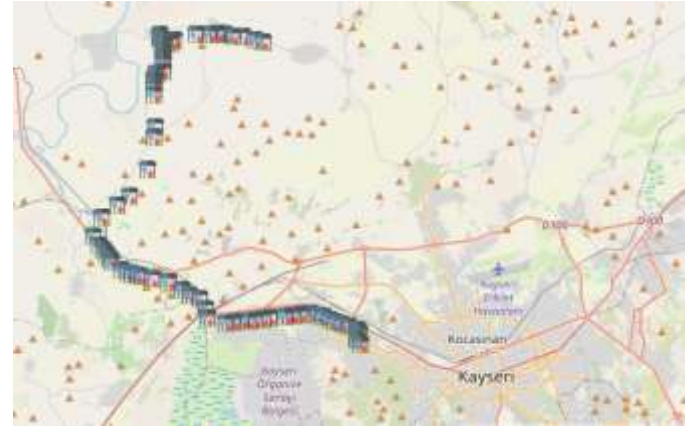


Fig.4. Selected bus route

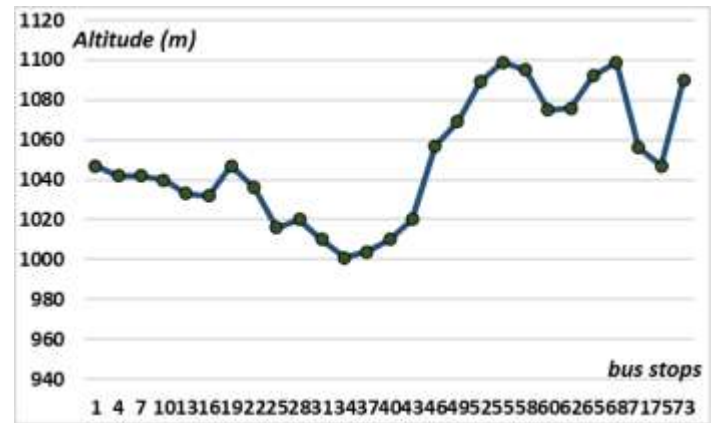


Fig.5. Topography for the R40 bus line. Source: Google Earth.

C. Calculations

This subsection presents the equations used to determine the concluding part; ERS costs, fees, savings, and estimates presented in a bulleted list. Then, the equations used to determine the cost of EB and ERS components are shown and concluded.

- ERS Costs

$$\sum TC_{ERS} = l_r \cdot c_{ers/km} \quad (1)$$

Where;

TC_{ERS} – total cost of implementing ERS (\$);

l_r – length of the road (km);

$c_{ers/km}$ – the cost per km (\$/km).

- Charge Cost

$$\sum TC_{el} = l_r \cdot c_{el/km} \quad (2)$$

Where;

TC_{el} – the total electricity cost (\$);

$c_{el/km}$ – the electricity cost per km (\$).

- User fee

$$U_{fee} = l_r^t \cdot u_{fee/km} \quad (3)$$

Where;

U_{fee} – the user fee (\$);

l_r^t – the total driven distance (km);

$u_{fee/km}$ – the cost per km of using the ERS (\$).

- Forecasts of diesel price

$$P_{in}^{diesel} = \frac{Pd_{2023}}{Pd_{2030}}^{\frac{1}{2030-2023}} \quad (4)$$

Where;

P_{in}^{diesel} – annual average increase of diesel in Turkey;

Pd_{2023} – diesel price in 2023;

Pd_{2030} – diesel price in 2030.

- Received and consumed energy

$$E_{cns} = t/60 \cdot E_{km} \cdot l_r^t \quad (5)$$

Where;

E_{cns} – the amount of electricity received;

t – spent time on ERS (minute);

E_{km} – energy use (kWh/km);

l_r^t – the total driven distance (km).

- Fuel Costs

$$C_f = l_r^d \cdot E_{km} \cdot c_f \cdot 365 \quad (6)$$

Where;

C_f – the annual cost for energy consumption;

l_r^d – the daily driven distance (km);

E_{km} – energy use (kWh/km);

c_f – fuel cost.

- Battery Cost

$$C_{bat} = S_{bat} \cdot c_{kW} \quad (7)$$

Where;

C_{bat} – cost of the battery (\$);

S_{bat} – battery capacity;

c_{kW} – cost of energy (\$/kWh).

- Energy receiver Cost

$$C_{rec} = c_{rec} \cdot n \quad (8)$$

Where;

C_{rec} – total cost of the receiver;

c_{rec} – cost for one receiver;

n – number of required receivers.

- Lifetime Cost

$$C_{lf} = C_{bus} + (n_y (C_{OM} + C_f)) \quad (9)$$

Where;

C_{lf} – lifetime cost;

C_{bus} – cost of the bus;

n_y – number of years;

C_{OM} – operation and maintenance cost;

C_f – fuel cost.

V. RESULTS

This section consists of two subsections: calculations and findings. The first part of this chapter shows the results of calculations using the formulas in subsection 3.4. The second

section presents the findings. In calculations, it is done to calculate the distance of the road to be electrified per kilometer. Some first comments:

1. The initial investment is equal to the construction cost of the ERS.

2. Annual savings are savings in operating costs of vehicles. Using EBs instead of fuel saves money. Both the annual savings and the annual cost are calculated according to the alternative fuel vehicle.

3. The total annual cost is the additional cost of buying an EB instead of a fuel bus. The additional cost for buses is assumed to be the same as the battery cost. It is assumed that an EB with a 120-kWh battery has the same price as a diesel bus. It's still a rather cautious assumption, given the estimates that an EB with a range of 70km would already be as cheap as fuel buses.

4. The second part of the annual cost is the maintenance cost of the ERS. It is accepted as 1% or 10.000\$/km of the annual investment cost (see Table 2).

In the sensitivity analysis, the effect of EBs needs to have a larger battery will be analyzed in order to make a correct calculation with an infrastructure alternative in the electrification of the road. This reduces the investment cost of the electric road while increasing the capital cost of the buses for the batteries. Boya [51] states that a 175-kWh battery will allow a range of 100 km. They state that the cost of such a battery was €60,000 in 2015. They estimate the cost will drop to €25,000 in 2025 and €19,000 in 2030. Therefore, this study assumes that the additional cost of upgrading the battery (120kWh) to cover 70 km will be €13,300 in 2030.

The expenses of adopting ERS in Kayseri city are discussed in this subsection, along with a cost comparison of the three ERS technologies—conductive rail, conductive overhead, and inductive. The bus route road is 35.5 km long and the cost of implementing ERS (one way) and maintenance cost are summarized in Table 6 varies between the different ERS technologies (see Figure 6).

TABLE VI
INSTALLATION AND MAINTENANCE COST FOR SELECTED BUS ROUTE

Type of ERS	ERS Cost (M\$)	Maintenance Cost (M\$)
Inductive	78,1	1,2
Conductive Overhead	49,7	0,6
Conductive Rail	28,4	0,4

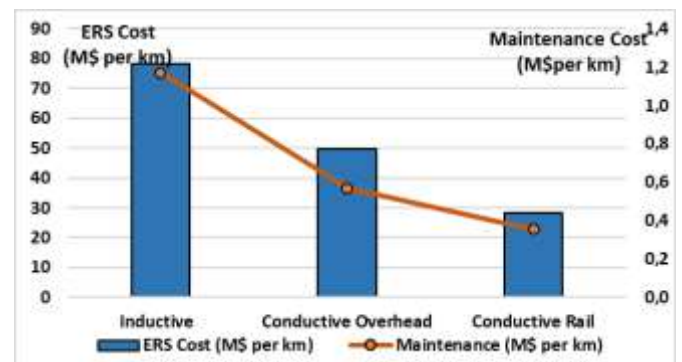


Fig.6. Total investment and maintenance cost to install 35.5 km of ERS.

A standard bus (12 meters) costs about 181.000\$ while an EB cost about 362.000\$ [52]. The maintenance cost varies between diesel and EBs. The maintenance cost of a diesel bus is about 15.000\$ per year and the maintenance cost of an EB is about 5000\$ per year. Table 7 compares for the cost of the ICE bus and EB.

TABLE VII
COMPARISON FOR COST OF ICE BUS AND EB

Bus Type	Bust Cost (\$)	The maintenance cost for 10 years (\$)	The fuel cost for 10 years (\$)	Total Cost (\$)
Diesel	81.000	150.000	1.422.734	1.753.734
Electric Bus	362.000	50.000	524.779	936.779

The total cost difference between EBs and diesel buses is 816.955\$. If society stops spending money at the individual level on bigger batteries and instead spends money collectively on electric road infrastructure, a small battery and an electrified road infrastructure can cost less than a larger long-range battery. If decision-makers stop spending money on bigger batteries and instead spend money on ERS infrastructure collectively, a small battery and electrified road infrastructure could cost less than a larger long-range battery. However, these figures are relatively small when compared to battery and bus costs; with an average of 2023 prices, it will be possible to establish ERS in Kayseri with a 10-year EB cost. Figure 7 shows the costs of the ICE bus and EB in ERS.

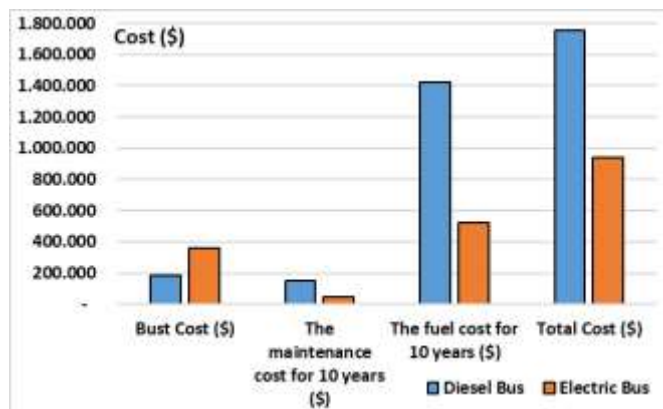


Fig.7. Cost comparison for diesel bus and EB

VI. CONCLUSION AND RECOMMENDATIONS FOR FURTHER WORK

This final section provides answers to the research questions. This section also includes discussions about limitations, research contributions, and some suggestions for future work.

The purpose of this work was to investigate whether roads can be developed and made smarter to promote greener. The main idea of electric road systems is to reduce CO₂ emissions, reduce the dependence of heavy vehicles on fossil fuels, provide good transportation for commercial purposes in the fossil-free society of the future, and at the same time

reduce vehicle costs by reducing the battery size needed in EVs.

In this study, the ERS scenario developed for the R40 bus line in the province of Kayseri is examined from the economic point of view of establishing an electric bus infrastructure with a range of 70 km. The cost of implementing this infrastructure is approximately \$28.4 million and the annual operating maintenance cost is approximately \$400,000. If a large battery is replaced with a smaller battery for each new bus sold, enough savings will be achieved after 10 years to electrify the 35.5km R40 bus line route in Kayseri.

Whether this study will benefit the future development of electrified road traffic is not difficult to answer at the time of writing this article, as this is the first article written for ERS for public transportation in Turkey. Despite the study's probable limitations, it has found that electrified road infrastructures are more economical under all assumptions than the corresponding battery-electric vehicle scenario. Additionally, smaller battery EVs are more affordable than larger battery EVs when combined with ERS. It can be concluded that by reducing battery deterioration and energy consumption, buses perform better when an ERS is installed.

If the ERS is to be expanded and used in Turkey in the future, government architecture, design, and design goals should also be followed. The government should also actively support the development of the design field in Turkey with exemplary actions from the public project. This must be done to strengthen the quality of the designed living environment. In addition, it is something that slows down the process a bit in our day of bureaucracy and decision-making. A design proposal that fits the technology's potential and sustainable intentions could accelerate the development of a fossil-free Turkey.

Whatever the choice of technology for power transmission, electricity routes to commercial operation will need payment systems to charge for infrastructure and energy use. This could be explored in future studies.

Placing solar panels on the road; asphalt road is unused land, instead of installing solar panels on the side of the road or directly on the road, solar panels can be placed on the carriageway. Driving on a shaded road has its advantages; tires provide a lower operating temperature and thus have a longer service life, the car's air conditioning needs to run less on hot days, and road repair costs are lower. This is also because the electricity required for ERS is obtained from solar panels.

REFERENCES

- [1] C. Yang, D. McCollum, R. McCarthy, and W. Leighty, "Meeting an 80% reduction in greenhouse gas emissions from transportation by 2050: A case study in California," *Transp. Res. Part D Transp. Environ.*, vol. 14, no. 3, pp. 147–156, May 2009, doi: 10.1016/j.trd.2008.11.010.
- [2] A. Ghosh, "Possibilities and Challenges for the Inclusion of the Electric Vehicle (EV) to Reduce the Carbon Footprint in the Transport Sector: A Review," *Energies*, vol. 13, no. 10, p. 2602, May 2020, doi: 10.3390/en13102602.

- [3] X. Yuan, L. Li, H. Gou, and T. Dong, "Energy and environmental impact of battery electric vehicle range in China," *Appl. Energy*, vol. 157, pp. 75–84, Nov. 2015, doi: 10.1016/j.apenergy.2015.08.001.
- [4] M.-K. Tran et al., "A Review of Range Extenders in Battery Electric Vehicles: Current Progress and Future Perspectives," *World Electr. Veh. J.*, vol. 12, no. 2, p. 54, Apr. 2021, doi: 10.3390/wevj12020054.
- [5] "Wireless electric road charges vehicles as they drive," *Smart city Sweden*, 2022. <https://smartcitysweden.com/best-practice/409/wireless-electric-road-charges-vehicles-as-they-drive/> (accessed Jan. 20, 2023).
- [6] M. Gaeta, C. Nsangwe Businge, and A. Gelmini, "Achieving Net Zero Emissions in Italy by 2050: Challenges and Opportunities," *Energies*, vol. 15, no. 1, p. 46, Dec. 2021, doi: 10.3390/en15010046.
- [7] W. Shoman, S. Karlsson, and S. Yeh, "Benefits of an Electric Road System for Battery Electric Vehicles," *World Electr. Veh. J.*, vol. 13, no. 11, p. 197, Oct. 2022, doi: 10.3390/wevj13110197.
- [8] H. H. Coban, A. Rehman, and A. Mohamed, "Analyzing the Societal Cost of Electric Roads Compared to Batteries and Oil for All Forms of Road Transport," *Energies*, vol. 15, no. 5, Mar. 2022, doi: 10.3390/en15051925.
- [9] H. H. Coban and W. Lewicki, "Daily Electricity Demand Assessment on The Example of The Turkish Road Transport System-A Case Study of The Development of Electromobility on Highways," *Transp. Geogr. Pap. PGS*, vol. 2022, no. 2, pp. 52–62, doi: 10.4467/2543859XPKG.22.011.16269.
- [10] S. Eslami, Y. Noorollahi, M. Marzband, and A. Anvari-Moghaddam, "District heating planning with focus on solar energy and heat pump using GIS and the supervised learning method: Case study of Gaziantep, Turkey," *Energy Convers. Manag.*, vol. 269, p. 116131, Oct. 2022, doi: 10.1016/j.enconman.2022.116131.
- [11] J. Schulte and H. Ny, "Electric road systems: Strategic stepping stone on the way towards sustainable freight transport?," *Sustain.*, vol. 10, no. 4, 2018, doi: 10.3390/su10041148.
- [12] D. Jelica, M. Taljegard, L. Thorson, and F. Johnsson, "Hourly electricity demand from an electric road system – A Swedish case study," *Appl. Energy*, vol. 228, no. June, pp. 141–148, 2018, doi: 10.1016/j.apenergy.2018.06.047.
- [13] S. Tongur and H. Sundelin, "The electric road system transition from a system to a system-of-systems," in 2016 Asian Conference on Energy, Power and Transportation Electrification (ACEPT), Oct. 2016, pp. 1–8, doi: 10.1109/ACEPT.2016.7811529.
- [14] R. C. Majhi, P. Ranjitkar, and M. Sheng, "Assessment of dynamic wireless charging based electric road system: A case study of Auckland motorway," *Sustain. Cities Soc.*, vol. 84, p. 104039, Sep. 2022, doi: 10.1016/j.scs.2022.104039.
- [15] R. Kühne, "Electric buses – An energy efficient urban transportation means," *Energy*, vol. 35, no. 12, pp. 4510–4513, Dec. 2010, doi: 10.1016/j.energy.2010.09.055.
- [16] M. Duan, G. Qi, W. Guan, C. Lu, and C. Gong, "Reforming mixed operation schedule for electric buses and traditional fuel buses by an optimal framework," *IET Intell. Transp. Syst.*, vol. 15, no. 10, pp. 1287–1303, Oct. 2021, doi: 10.1049/itr2.12098.
- [17] B. Verbrugge, M. M. Hasan, H. Rasool, T. Geury, M. El Baghdadi, and O. Hegazy, "Smart Integration of Electric Buses in Cities: A Technological Review," *Sustainability*, vol. 13, no. 21, p. 12189, Nov. 2021, doi: 10.3390/su132112189.
- [18] Z. Bi, L. Song, R. De Kleine, C. C. Mi, and G. A. Keoleian, "Plug-in vs. wireless charging: Life cycle energy and greenhouse gas emissions for an electric bus system," *Appl. Energy*, vol. 146, pp. 11–19, May 2015, doi: 10.1016/j.apenergy.2015.02.031.
- [19] B. Csonka, "Optimization of Static and Dynamic Charging Infrastructure for Electric Buses," *Energies*, vol. 14, no. 12, p. 3516, Jun. 2021, doi: 10.3390/en14123516.
- [20] N. Jones, "More than a billion cars to hit the road," *Nature*, Nov. 2006, doi: 10.1038/news061113-15.
- [21] H. Sundelin, M. G. H. Gustavsson, and S. Tongur, "The maturity of electric road systems," 2016 Int. Conf. Electr. Syst. Aircraft, Railw. Sh. Propuls. Road Veh. Int. Transp. Electr. Conf. ESARS-ITEC 2016, 2016, doi: 10.1109/ESARS-ITEC.2016.7841380.
- [22] H. Alfredsson and E. von Essen, "Cost optimized electrification of Lund city bus traffic using Elonroad electric road system," *Lund University*, 2018.
- [23] "Roadmap for an overhead catenary system for trucks," 2022. <https://www.ifeu.de/en/project/roadmap-for-an-overhead-catenary-system-for-trucks/> (accessed Jan. 20, 2023).
- [24] "International Research Collaboration on Electric Road Systems II (CollERS2)," *Electric-road-systems*, 2022. <https://electric-road-systems.eu/e-r-systems/> (accessed Jan. 20, 2023).
- [25] Sundelin Håkan, "Swedish-German research collaboration on Electric Road Systems," Heidelberg, Mar. 2021. Accessed: Jan. 20, 2023. [Online]. Available: <https://policycommons.net/artifacts/1966846/swedish-german-research-collaboration-on-electric-road-systems/2718611/>.
- [26] E. Berlin et al., "StratON Bewertung und Einführungsstrategien für oberleitungsgebundene schwere Nutzfahrzeuge," 2020. Accessed: Jan. 20, 2023. [Online]. Available: <https://www.oeko.de/fileadmin/oekodoc/StratON-O-Lkw-Endbericht.pdf>.
- [27] Sarah Wray, "Tel Aviv pilots 'electric road' to charge buses," *Cities-today*, Aug. 23, 2020. <https://cities-today.com/tel-aviv-pilots-electric-road-to-charge-buses/> (accessed Jan. 20, 2023).
- [28] Bernhardt Jens, "Electric road for ebus charging on test in Lund/Sweden," Nov. 28, 2019.
- [29] Y. J. Jang, "Survey of the operation and system study on wireless charging electric vehicle systems," *Transp. Res. Part C Emerg. Technol.*, vol. 95, pp. 844–866, Oct. 2018, doi: 10.1016/j.trc.2018.04.006.
- [30] "Siemens demonstrates first eHighway system in the U.S.," *Siemens*, Nov. 08, 2017AD. <https://press.siemens.com/global/en/pressrelease/siemens-demonstrates-first-ehighway-system-us> (accessed Jan. 20, 2023).
- [31] Rise Torun and Nørbech Tom, "ELinGO - Electrification of heavy freight transport," *Sintef*, 2022. <https://www.sintef.no/projectweb/elingo/english/> (accessed Jan. 20, 2023).
- [32] Powerelectronicsnews, "Electric Bus Utilizing Wireless Electric Road," Oct. 20, 2021. <https://www.powerelectronicsnews.com/electric-bus-utilizing-wireless-electric-road/> (accessed Jan. 20, 2023).
- [33] Otilia Drăgan, "An e-Bus Starts Testing World's First Wireless Electric Road, in Breakthrough Project," *Autoevolution*, Oct. 20, 2021. <https://www.autoevolution.com/news/an-e-bus-starts-testing-worlds-first-wireless-electric-road-in-breakthrough-project-172164.html> (accessed Jan. 20, 2023).
- [34] H. Basma, C. Mansour, M. Haddad, M. Nemer, and P. Stabat, "Energy consumption and battery sizing for different types of electric bus service," *Energy*, vol. 239, p. 122454, Jan. 2022, doi: 10.1016/j.energy.2021.122454.
- [35] "Leading the Wireless WAVE," 2022. <https://waveipt.com/> (accessed Jan. 20, 2023).
- [36] "Wireless electric charging infrastructure for German roads," *Electreon*, 2023. <https://electreon.com/projects/karlsruhe> (accessed Jan. 20, 2023).
- [37] Nepori Andrea, "Italian highway recharges electric vehicles as they go," *Domus*, Jun. 21, 2022.
- [38] Jitchotvisut Janaki, "Italy's Inductive Highway Project Going Very Well, Says Stellantis," *Motor1*, Jun. 20, 2022.
- [39] "Fabric Project Paves the Way for Large Scale Deployment of Electromobility," *Ertico*, Mar. 17, 2014. <https://erticonetwork.com/fabric-project-paves-the-way-for-large-scale-deployment-of-electromobility/> (accessed Jan. 20, 2023).
- [40] "e-Road Arlanda, Stockholm," *Elways*, 2022. <https://elways.se/elways/projects/> (accessed Jan. 20, 2023).
- [41] "Developing a functional viable solution for dynamic charging," *Electreon*, 2022. <https://electreon.com/projects/echarge-bast> (accessed Jan. 20, 2023).
- [42] "eHighway – Electrification of road freight transport," *Siemens*, 2022. <https://www.mobility.siemens.com/global/en/portfolio/road/ehighway.html> (accessed Jan. 20, 2023).
- [43] S. Lee, J. Huh, C. Park, N. S. Choi, G. H. Cho, and C. T. Rim, "On-Line Electric Vehicle using inductive power transfer system," 2010 IEEE Energy Convers. Congr. Expo. ECCE 2010 - Proc., pp. 1598–1601, 2010, doi: 10.1109/ECCE.2010.5618092.

- [44] "Turkey, National Inventory Report (NIR)," Apr. 2022. Accessed: Jan. 20, 2023. [Online]. Available: <https://unfccc.int/documents/461926>.
- [45] Temizer Murat, "Turkey's first electric bus hits the roads," Istanbul, Dec. 2014. Accessed: Jan. 20, 2023. [Online]. Available: <https://www.aa.com.tr/en/energy/regulation/turkeys-first-electric-bus-hits-the-roads/12483>.
- [46] "Turkey's First Ultra-Charging Electric Buses Started Carrying Passengers," Samsun, Aug. 2022. Accessed: Jan. 20, 2023. [Online]. Available: <https://www.raillynews.com/2022/08/Turkey%27s-first-ultra-chargeable-electric-buses-started-to-carry-passengers/>.
- [47] A. S. Al-Ogaili et al., "Review of the Estimation Methods of Energy Consumption for Battery Electric Buses," *Energies*, vol. 14, no. 22, p. 7578, Nov. 2021, doi: 10.3390/en14227578.
- [48] J. A. Manzoli, J. P. Trovão, and C. H. Antunes, "A review of electric bus vehicles research topics – Methods and trends," *Renew. Sustain. Energy Rev.*, vol. 159, p. 112211, May 2022, doi: 10.1016/j.rser.2022.112211.
- [49] "Kayseri Ulasim," 2023, Accessed: Jan. 20, 2023. [Online]. Available: <https://www.kayseriulasim.com/tr/FaaliyetAlanlarimiz/otobus-isletmesi/hakkinda#:~:text=Kayseri Ulaşım A.Ş'nin,yönetiminde toplam 613 otobüs mevcuttur>.
- [50] G. Berckmans, M. Messagie, J. Smekens, N. Omar, and L. Vanhaverbeke, "Cost Projection of State of the Art Lithium-Ion Batteries for Electric Vehicles Up to 2030," *Energies*, vol. 10, no. 9, p. 1314, Sep. 2017, doi: 10.3390/en10091314.
- [51] B. Zhou et al., "Real-world performance of battery electric buses and their life-cycle benefits with respect to energy consumption and carbon dioxide emissions," *Energy*, vol. 96, pp. 603–613, Feb. 2016, doi: 10.1016/j.energy.2015.12.041.
- [52] Akgunduz Bilal, "Yerli Elektrikli Otobüs Modelleri, Teknik Özellikleri," Feb. 02, 2022. <https://www.dolubatarya.com/haber/yerli-elektrikli-otobus-modelleri-teknik-ozellikleri> (accessed Jan. 20, 2023).

BIOGRAPHY



HASAN HUSEYİN COBAN born in 1985, obtained his BSc degree in electronics engineering from Vilnius Gediminas Technical University (VGTU) in 2008. He received the MSc. diploma in Electrical Engineering from Klaipeda University in 2010, and PhD degrees in Riga Technical University in 2016. He worked as

a project engineer in the industry between 2010 and 2018. His research interests are Sustainable Energy Technologies, Renewable energy systems, Energy Policy, Energy Access, ANN, Electric Vehicles. Since Feb. 2018, he is an Assistant Professor at Ardahan University Electrical & Electronics Engineering department. He teaches Power Systems, Energy Systems, and Optimization for Engineers courses.

Performance Evaluation of Jaccard-Dice Coefficient on Building Segmentation from High Resolution Satellite Images

İsa Ataş

Abstract— In remote sensing applications, segmentation of input satellite images according to semantic information and estimating the semantic category of each pixel from a given set of tags have of great importance for the automatic tracking task. It is important in situations such as building detection from high resolution satellite images, city planning, environmental preparation, disaster management. Buildings in metropolitan areas are crowded and messy, so high-resolution images from satellites need to be automated to detect buildings. Segmentation of remote sensing images with deep learning technology has been a widely considered area of research. The Fully Convolutional Network (FCN) model, a popular segmentation model, is used for building detection based on pixel-level satellite images. With the U-Net model, which was developed for biomedical image segmentation and modified in our study, performance analysis was performed for building segmentation from satellite images using customized loss functions such as Dice Coefficient and Jaccard Index measurements. Dice Coefficient score was obtained 84% and Jaccard Index score was obtained 70%. In addition, the Dice Coefficient score increased from 84% to 87% by using the Batch Normalization (BN) method instead of the Dropout method in the model.

Index Terms—Convolutional neural network, remote sensing imagery, semantic segmentation, U-Net.


I. INTRODUCTION

Buildings are identified as an important object class in remote sensing images. Identifying the spatial location and shape of buildings in these images takes an important role in assessment issues such as geographic planning, mapping, and post-disaster reconstruction. In particular, determining the number of buildings remaining in natural disasters such as earthquake events and subtract the degree of damage are

important indicators. Traditional remote sensing image classification methods use spectral, and spatial correlation information of the desired object to determine the class content of ground objects. Multi-scale feature aggregation etc. it is aimed to extract the spectral and shape properties of buildings by using techniques such as [1]. With the rapid development of deep learning recently, it has made great progress in the fields of agriculture [2-3], health [4], robotics and security [5], natural language processing [6], as well as in the fields of remote sensing, semantic segmentation and target recognition [1]. Deep learning methods are in the direction of producing more stable results for image recognition in remote sensing with high accuracy and high computational speeds [1]. With the development of remote sensing technology, many researchers were tried to develop, try and explore new methods to increase the accuracy and the speed of automatic classification algorithms from satellite images. In image processing applications, image segmentation and labeling are a very demanding task. Convolutional Neural Network (CNN) is a more effective approach than traditional image segmentation methods. CNN automatically learns multi-level representations that map the original input to specified binary or multiple labels or sequential vectors. CNN's powerful "representational learning" capability greatly simplifies the feature design upon building detection and has shown promising results [7]. U-Net is structurally one of the simplified CNN architectures used for image segmentation. With the semantic segmentation method, studies are carried out on many complex tasks. Autonomous vehicles [8], biomedical image diagnostics [9], geo-sensing [10], and precision agriculture [11] are a few of them.

In this study, building extraction was performed from high resolution images with Semantic Segmentation using FCN. The contribution of study to previous research is to investigate which resolution might be more useful by comparing object extraction and segmentation approaches from different resolution satellite images. In addition, the effect of the dropout layer and batch normalization layers on the U-Net architecture was analyzed. F-measure (Dice Coefficient) and Intersection Unit (IoU) / (Jaccard Coefficient) metrics were used to distinguish the building class from satellite images and to quantitatively evaluate the performance of the segmentation method. The remainder of the article was organized as follows. In the second part, references were made to previous studies in the literature. In the third chapter, the materials and methods used were explained. In the fourth section, the

İSA ATAŞ, is with Computer Technologies Department, Diyarbakır Vocational School of Technical Sciences, Dicle University, Diyarbakır, Turkey, (e-mail: isa_atas@dicle.edu.tr).

 <https://orcid.org/0000-0003-4094-9598>

Manuscript received December 1, 2022; accepted Jan 16, 2023
DOI: [10.17694/bajece.1212563](https://doi.org/10.17694/bajece.1212563)

analyzes and result graphics were shown. The existing study are summarized in the conclusion.

II. RELATED WORK

Wu et al. Firstly, it skillfully combined the spectral features and shape features of buildings to distinguish buildings from remote sensing images using the traditional remote sensing image feature extraction method [12]. Huang et al. used the multi-feature fusion method to extract features of buildings in remote sensing images and finally chose to use the SVM classifier to classify ground objects [13]. Xin et al. used the spatial position relationship between shadows and buildings in densely built areas and used the graph cutting algorithm to accurately draw the outlines of buildings [14]. Acar et al. analyzed Radarsat-2 satellite images to retrieve soil moisture using semi-empirical models and machine learning methods. By extracting the feature vector from preprocessed SAR image pixels, they stated that machine learning techniques outperform semi-empirical models [15]. Xu-dong et al. it has achieved good results using the pixel-based classification technique to process the image using the object-oriented method [16]. After machine learning techniques became widespread, Deep CNN (DCNN) started to be used in current image classification and extraction. Wu et al. performed end-to-end building segmentation from aerial images using multi-constraint FCN architecture [17]. Yuan, a simple FCN model architecture was proposed that combines several activation layers in pixel level estimation. In addition, the marked distance function of building boundaries with enhanced representativeness was introduced to present the outputs [18]. Chen et al. applied various cutting-edge deep U-Net frameworks such as pyramid scene decomposition network, feature pyramid network (FPN) and multi-scale feature fusion FPN to generate roof detection from large-scale comparative aerial images [19].

III. MATERIALS AND METHODS

A. Dataset

A dataset consisting of .tif files containing real aerial photographs collected from the satellite and masked images were used. The input images of the proposed model are in RGB format. The dataset consists of a total of 408 images and tag masks with a resolution of 512×512 (ranging from 0.3 to 2.5 m) received from the satellite, prepared by the Wuhan University Photogrammetry and Computer Vision (GPCV) Group [20].

B. Preprocessing

Images with few buildings, no buildings and poor resolution were removed from the dataset during the process. After the elimination, only a total of 360 images could be obtained. Since it would be insufficient to train and test this number of image models, data augmentation methods were applied to the images. To increase sample images, images and mask images were augmented using the data augmentation method rotation (90 degrees right) and (90 degrees left). With this technique, 2 additional variations of each original image were created. As a

result of this process, the dataset was tripled and 1080 images and related masked images were prepared. We resized the image and tag masks at 256×256 , 128×128 , and 64×64 pixels along with the existing 512×512 -pixel dataset to research on, train and test the optimal patch size on the FCN network. By comparing the performance of images and masks in four different pixels in the U-Net model, we observed that the 256×256 size was more suitable than the other sizes. From these prepared images, 972 images and corresponding labeled images ($\approx 80\%$) were randomly selected to be used as training set, and the remaining 108 images and corresponding labeled images ($\approx 20\%$) were used as test dataset.

Here, by applying random cropping from the image data, we cropped large images into small patches in the desired pixels and obtained more data for our model. Similarly, we applied the same for the masked dataset. In our preprocessing, we visualized both cropped patches (image and mask) and checked whether the newly formed image and tagged masks were aligned correctly. Since the tags we extracted from the satellite images consisted only of the "buildings" layer, we converted the tag images into binary masks of four different sizes. In the labels, we pixelated the buildings as 1 and the outside of the building as 0. In Fig. 1 shows the sample satellite image from the dataset, as well as the 256×256 patch image and mask that we referenced in our study and formed after cropping.

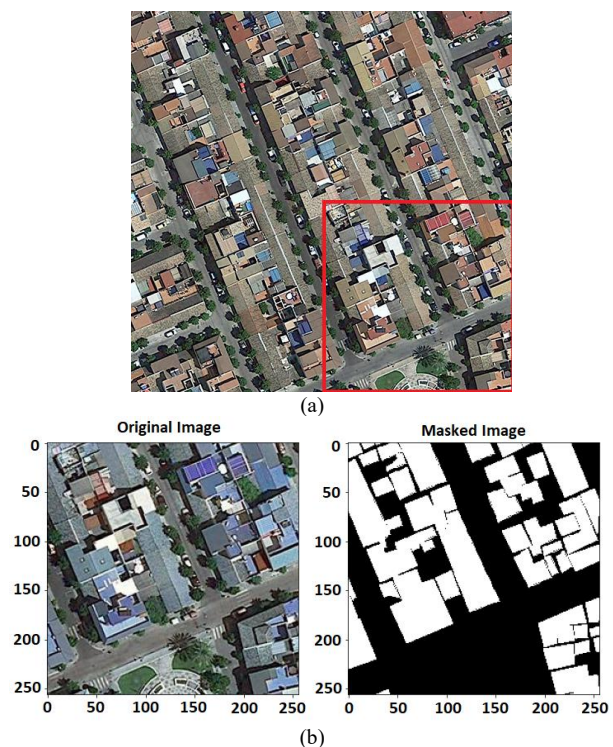


Fig.1. (a) Satellite image (b) Cropped Satellite image and corresponding mask with buildings identified in white.

C. Architecture of the model (U-Net)

The U-Net architecture was developed by Olaf Ronneberger et al. [21] for Biomedical Image Segmentation. It aims at image segmentation using a small training dataset. The general

structure of U-Net, which is a typical encoder-decoder architecture and modified in this study, is shown in Fig. 2.

U-Net Architecture includes two paths, convolution (Contraction path) and deconvolution (Expanding path) layers. The first way is the narrowing path, also called an encoder, which captures the input image at multiple different levels of feature representations. The encoder is just a traditional stack of convolution and max pooling layers. The second path is the symmetric expansion path, also called a decoder, which is to semantically project the distinctive features (lower resolution) into the pixel space (higher resolution). Basically, U-Net is built on FCN [22]. For this reason, it is also referred to as end-to-end FCN. It contains only Convolutional layers and no dense layers because it can accept images of any size. Feature maps created by convolution layers are clipped to input size. Clipping deals with the loss of edge pixels in the convolution process and subtracts the size of the convolution results, consistent with the deconvolution results. In the U-Net architecture, jump links (copy and crop) are added between the narrowing path and the widening path, allowing precise localization as well as features. Clipped feature maps are combined with deconvolution results via shortcut links. The network applies 1×1 convolution to the feature map to label the pixels and produces the segmentation result.

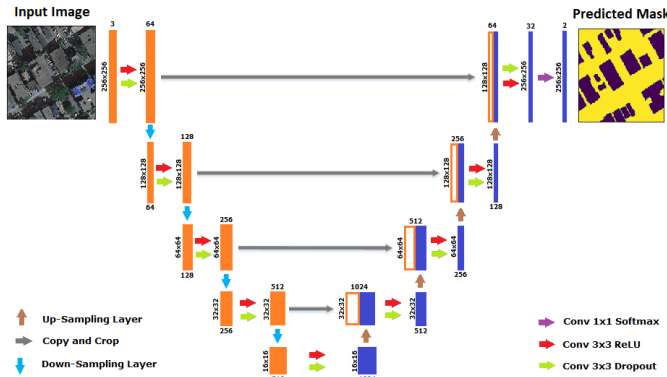


Fig. 2. Architecture of U-Net

In the U-Net architecture model, the orange boxes represent cross sections of the feature maps. The dimensions of each feature map are indicated on the left side, and the number of channels is indicated above. White boxes represent half of the channels with feature maps from which they were copied. Arrows represent actions indicated by the indicator; gray arrows represent copying (skipping links). The purpose of the encoder layer in the model is to extract feature information from the images until the model is at the bottom. The purpose of the decoder layer is to find the required area that the model should represent from the fragmentary patterns.

In our study, we modified the U-Net architecture for the building segmentation problem. We trained the model from scratch without using pre-trained weights with biomedical images. First, we replaced the stochastic gradient descent algorithm (SGD) [23] with the Adam [24] optimizing algorithm, which is known to converge faster during training. Next, we changed the dimensions of the input images to 256×256 , since the U-net architecture was designed for images with a size of 572×572 . In our experimentally applied model, the

optimum learning rate was chosen as 0.0001. To reduce the computational cost, we chose to simplify the model by changing the filter numbers of the convolutional layers to 16, 32, 64, and 128 from 64, 128, 256, and 512 in the U-Net architecture. In end-to-end training, we added “dropout” to avoid overfitting when training the model in the decoder section after rectified linear unit (ReLU) activation in all convolution layers.

Due to the limited amount of GPU memory, we set the training batch size to 16 and maximum number of periods according to Epoch 100 in this study. We got the output image from the model by adding CNN layer with SoftMax instead of sigmoid as activation at the end of the expanding path layer. As a result, our app is based on Keras [25] using a TensorFlow backend. All training and testing processes were performed in Google Colab environment [26] with 13,342 RAM - Tesla K80 GPU - NVIDIA T4 GPUs Card.

D. Metrics

We used precision, recall, F-measure (Dice Coefficient) and Intersection of Unity (IoU) / (Jaccard Coefficient) metrics to distinguish building class from satellite images and to quantitatively evaluate the performance of the segmentation method. Segmentation based on Dice and Jaccard coefficients is widely used for image segmentation as it allows to deal with class imbalance [27]. Precision and recall are defined using true positives (TP), false positives (FP), and false negatives (FN).

1) Precision

The precision refers to the ratio of positive data identified as true to all data identified as true. It is calculated by following equation [28].

$$\text{Precision} = \text{TP} / (\text{TP} + \text{FP}) \quad (1)$$

2) Recall

The recall refers to the ratio of positive data identified as true to the sum of positive data identified as true and negative data identified as false. It is calculated by following equation [28].

$$\text{Recall} = \text{TP} / (\text{TP} + \text{FN}) \quad (2)$$

3) Jaccard Index

The first of the metrics used to evaluate the education score in our study is the Jaccard Index, also known as IOU (Intersection over Union). The Jaccard index, also known as the Jaccard similarity coefficient, is a statistic used to measure similarity. It was explanation by Grove Karl Gilbert in 1884 as the verification rate [29]. It was later developed independently by Paul Jaccard [30]. The Jaccard coefficient measures similarity between finite sets of samples and is defined as the size of the intersection divided by the size of the union of sample sets. It is calculated by following equation. $0 \leq J(A, B) \leq 1$

$$J(A,B)=|A \cap B|/(|A|+|B|-|A \cap B|)$$

$$J(A,B) = (\text{precision} \times \text{recall}) / (\text{precision} + \text{recall} - (\text{precision} \times \text{recall})) \quad (3)$$

4) Dice Coefficient

The second of the metrics used to evaluate the education score in our study is the Dice Coefficient, also known as the F1 score. Independently developed by Thorvald-Sørensen [31] and Lee-Raymond Dice [32], the Sørensen-Dice coefficient is a statistic used to measure the similarity of two samples. It is calculated by following equation.

$$0 \leq J(A, B) \leq 1$$

$$D(A,B) = 2 \times (|A \cap B|) / (|A| + |B|)$$

$$D(A,B) = 2 \times (\text{precision} \times \text{recall}) / (\text{precision} + \text{recall}) \quad (4)$$

IV. EXPERIMENTS/RESULTS/DISCUSSION

In order to obtain a successful U-Net training model in the segmentation task, some hyperparameters have been modified. In addition, analyzes were made to obtain loss and metric graph curves. Initially, training and validation analyzes of four different datasets with 64×64 , 128×128 , 256×256 , and 512×512 resolutions were made and the loss graphs are shown in Fig. 3. Considering the training and validation graph curves, convergences were observed differently in four different patch sizes. It is seen in the graphs that as the patch size increases, convergence tends to improve and then deteriorate.

Peak value was obtained with 256×256 pixels size providing the best IoU (Jaccard Coef.) 84.70%. The IoU plots of the datasets according to the four different patch sizes are shown in Fig. 4. Generally, in theory larger patches provide more information and accurate predictions. However, fine resolution images require a longer training time or added feature information may be unnecessary and affect the model training process [33]. However, cropped patches with a small window size may contain information that does not have sufficient properties for the model and may increase the risk of overfitting. Also, small size patches cannot complete all down sampling of the U-Net model [34].

The 64×64 satellite images failed to predict the building class with the modified U-Net model and the model was unable to complete the subsamples. It has followed an unstable path as it has not been able to extract adequate feature maps. It has been observed that the model building images from 128×128 and 512×512 sized satellite images have difficulty in separating from other objects. Therefore, in this study, a patch size of 256×256 pixels was chosen, which is suitable for the U-Net model. Intuitive comparisons of image segmentation results of input satellite images of 4 different sizes in model training are shown in Fig. 5.

The performances of the dice coefficient and jacquard index, which are among the missing metrics, on the model were analyzed.

The comparison of actual image, mask and predicted mask in Dice Coefficient and Jaccard-Index mode is shown in Fig. 6.

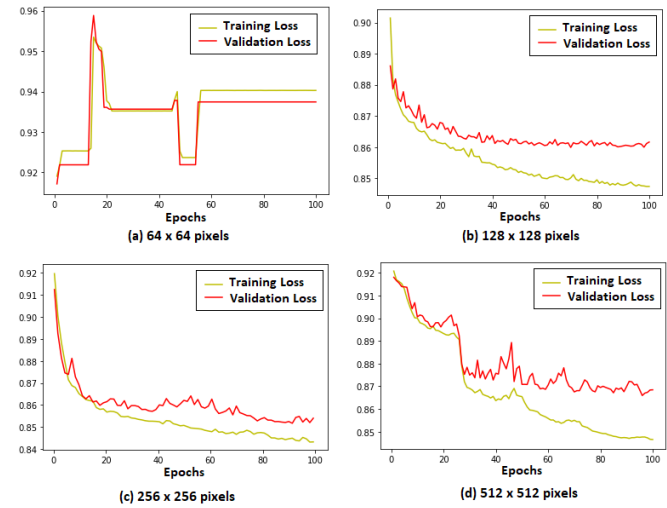


Fig. 3. Loss values of the proposed model over 100 epochs

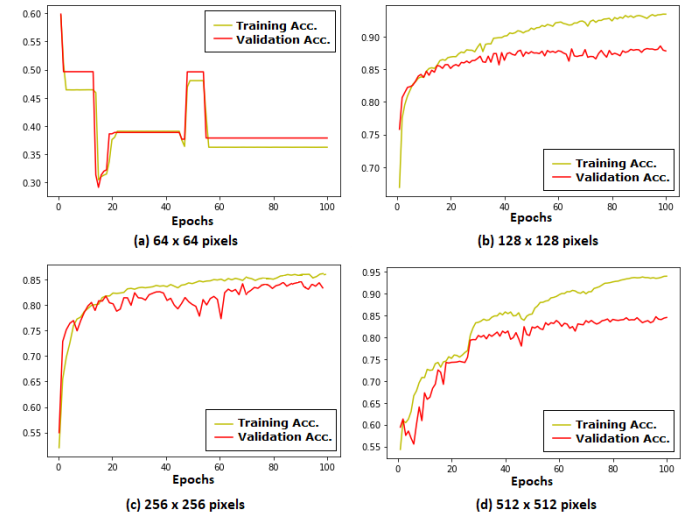


Fig. 4. Accuracy values of the proposed model over 100 epochs

Dice coefficient and Jaccard index loss curves on the training and validation dataset are shown in Fig. 7. The metric curves of the Dice coefficient and the Jaccard index are shown in Fig. 8 comparatively. When we look at the performances of both metrics, it is seen that the Dice coefficient is more consistent than the Jaccard index, and it provides superiority in loss and similarity rates. It also has realized less convergence in the training phase. In terms of improving the performance of the modified model, Dropout and BN methods were tried and analyzed as the threshold value before ReLU activation in the contraction path layer. The performances of the Dropout and BN index, which are among the missing metrics, on the model were analyzed. The comparison of actual image, mask and predicted mask in Dropout and BN mode is shown in Fig. 9. Dropout and BN loss curves on the training and validation dataset are shown in Fig. 10.

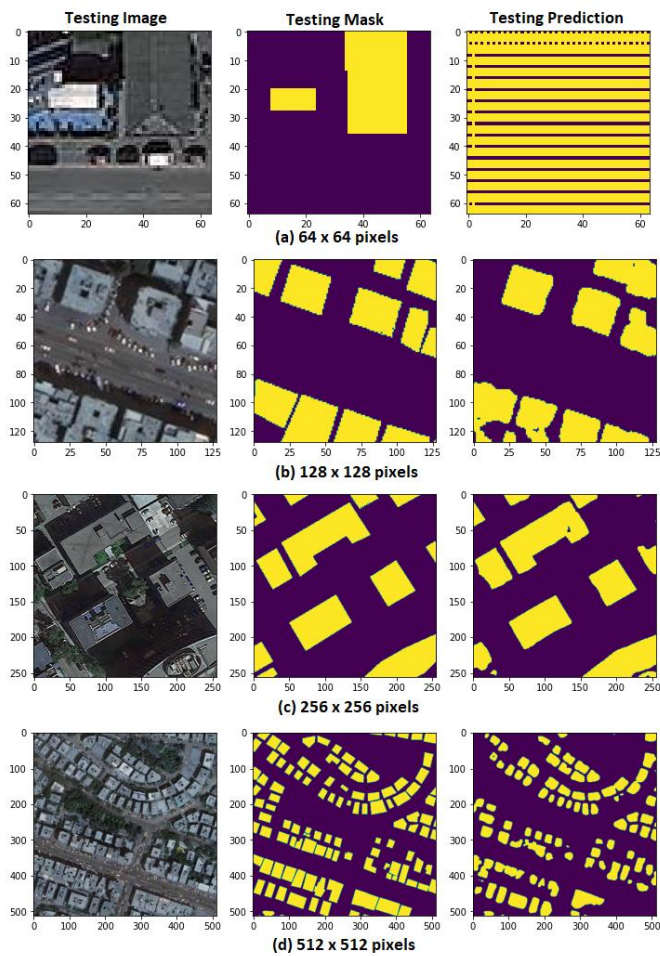


Fig. 5. Satellite images of four different dimensions, ground reality masks and masks predicted

Looking at the loss and metric graphs, it is seen that BN improves model performance and follows a stable path compared to the dropout method.

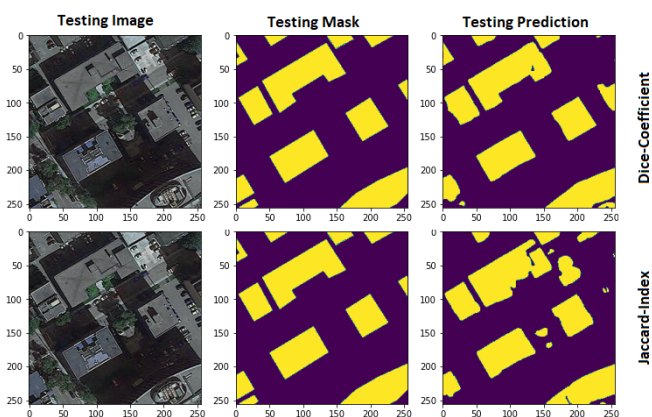


Fig. 6. Comparison of real image, mask and predicted mask in Dice-Coefficient and Jaccard-Index mode

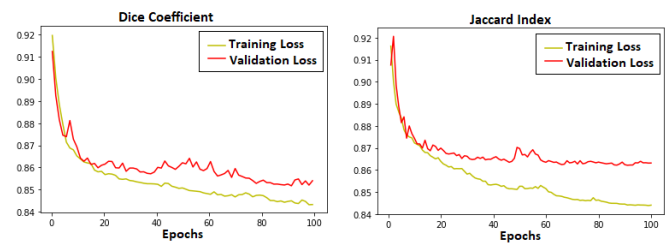


Fig. 7. Training and validation dataset: loss curves of Dice coefficient and Jaccard index

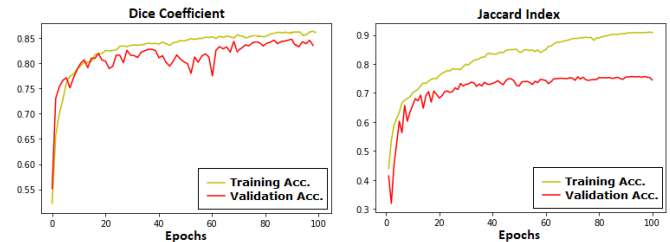


Fig. 8. Training and validation dataset: accuracy curves of Dice coefficient and Jaccard index

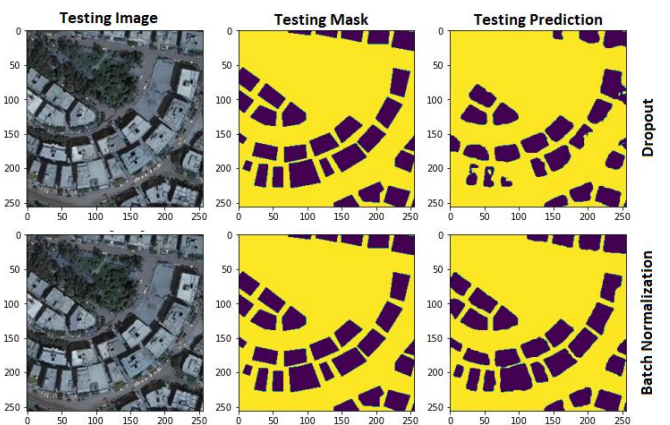


Fig. 9. Comparison of real image, mask and predicted mask in Dropout and Batch Normalization mode

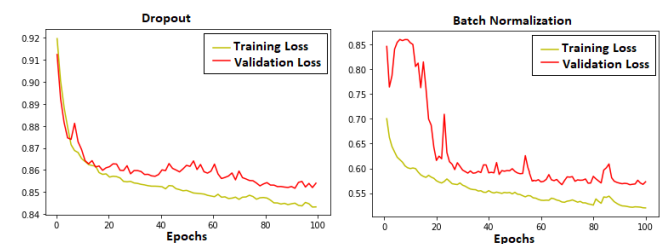


Fig. 10. Training and validation dataset: loss curves of Dropout and Batch Normalization

The metric curves of the Dropout and BN are shown in Fig. 11 comparatively. While the loss rate is 0.84 in the dropout method, it is seen that the loss rate decreases to 0.20 when the BN method is used. In addition, it was observed that the Dice-Coefficient metric value increased from 0.83 to 0.87 when the BN method was used instead of the dropout method.

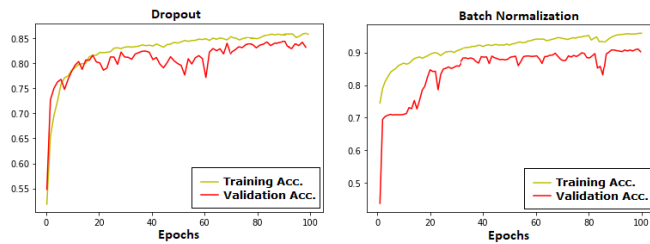


Fig. 11. Training and validation dataset: accuracy curves of Dropout and Batch Normalization

A rational comparison could not be made because experimental studies were conducted on different satellite image datasets and different metrics were used in the literature. However, a summary of the quantitative results based on similar studies is given in Table 1.

TABLE I
QUANTITATIVE COMPARISON OF SATELLITE IMAGE
SEGMENTATION STUDIES

Reference	Dataset	Model	Jaccard %	Dice %
[35]	Massachusetts	GMEDN	70.39	-
[36]	WHU	ESFNet	85.34	-
[37]	Carvana	U-Net	-	68.70
[38]	Inria	SegNet	70.14	-
[39]	MBRSC	U-Net	-	87.00
Our study	GPCV	U-Net	70.27	87.13

V.CONCLUSION

In this study, an end-to-end approach is adopted for binary mask classification using a small number of satellite image data. The prepared dataset was applied and evaluated in the CNN-based U-Net architecture developed by Ronneberger et al. The Keras framework and Python language were used to implement the U-Net model, facilitate data augmentation, and increase the robustness of the training. A patch size of 256×256 pixels was selected as an appropriate window size for the U-Net model in our apps. The model tested for the 64×64 pixel patch showed poor performance and resulted in weak compatibility. Since the test could not make meaningful predictions on the data, it was observed that high loss occurred with low accuracy. On the other hand, in the analysis of 128×128 and 512×512 pixel patches, overfitting occurred with poor performance. F-measure (Dice Coefficient) and Intersection Unit (IoU) / (Jaccard Coefficient) metrics were used to quantitatively evaluate and compare the performance of the proposed method. In comparison, it was seen that the dice coefficient was more consistent than the Jaccard index and was superior in loss and similarity rates. The BN process, in which the biased output distribution obtained from the previous layer is corrected, and the dropout process, in which some of the feature information obtained from the previous layer are randomly disabled are compared to observe the effect on the model. Dice Coefficient score was obtained 84% and Jaccard Index score was obtained 70%. In addition, the Dice Coefficient accuracy score increased from 84% to 87% by using the Batch Normalization (BN) method instead of the Dropout method in the model. As a result of the analysis, it

was determined that BN has a noticeable effect on performance compared to the Dropout layer.

In conclusion, considering the qualitative and quantitative results obtained in the analyzes performed in this study, the modified U-Net model performed well overall with its small dataset size and limited computational capacity.

REFERENCES

- [1] Q. Han, Q. Yin, X. Zheng, Z. Chen, "Remote sensing image building detection method based on Mask R-CNN." *Complex & Intelligent Systems*, 8(3), 1847-1855, 2022.
- [2] M. Ataş, "Fıstık sınıflandırma sistemi için Siirt fıstığı imgelerinden gürbüz özneteliklerin çıkarılması." *Dicle Üniversitesi Mühendislik Fakültesi Mühendislik Dergisi* 7(1):93-102, 2016.
- [3] E. Acar, "Detection of unregistered electric distribution transformers in agricultural fields with the aid of Sentinel-1 SAR images by machine learning approaches." *Computers and Electronics in Agriculture*, 175, 105559, 2020.
- [4] A. D. Yetis, M. I. Yesilnacar, M. Atas, "A machine learning approach to dental fluorosis classification." *Arabian Journal of Geosciences*, 14(2):1-12, 2021.
- [5] M. Atas, Y. Dogan, İ. Atas, "Chess playing robotic arm." In 2014 22nd Signal Processing and Communications Applications Conference (SIU) (pp. 1171-1174). IEEE, 2014.
- [6] C. Özdemir, M. Ataş, A. B. Özer, "Classification of Turkish spam e-mails with artificial immune system." 21st Signal Processing and Communications Applications Conference (SIU). IEEE, 2013.
- [7] S. Ji, S. Wei, M. Lu, "Fully convolutional networks for multisource building extraction from an open aerial and satellite imagery dataset." *IEEE Transactions on Geoscience and Remote Sensing*, 57(1), 574-586, 2018.
- [8] Ç. Kaymak, A. Uçar, "Semantic Image Segmentation for Autonomous Driving Using Fully Convolutional Networks." *International Artificial Intelligence and Data Processing Symposium (IDAP)*, 2019. DOI: 10.1109/IDAP.2019.8875923.
- [9] A. Valizadeh, M. Shariatee, "The Progress of Medical Image Semantic Segmentation Methods for Application in COVID-19 Detection." *Comput Intell Neurosci*. 2021, DOI: 10.1155/2021/7265644.
- [10] A. Mousavian, J. Kosecka, "Semantic Image Based Geolocation Given a Map." DOI: 10.48550/arXiv.1609.00278.
- [11] T. Anand, S. Sinha, M. Mandal, V. Chamola, F. R. Yu, "AgriSegNet: Deep aerial semantic segmentation framework for IoT-assisted precision agriculture." *IEEE Sensors Journal*, 21(16), 17581-17590, 2021.
- [12] W. Wu et al., "Building extraction from high resolution remote sensing imagery based on spatial-spectral method." *Geomat Inf Sci Wuhan Univ* 7:800-805, 2012.
- [13] X. Huang et al., "Classification of high spatial resolution remotely sensed imagery based upon fusion of multiscale features and SVM." *J Remote Sens* 11:48-54, 2007.
- [14] F. Xin, C. Shanxiong, "High-resolution remote sensing image building extraction in dense urban areas." *Bull Surv Mapp*, 2019.
- [15] H. Acar, M. S. Özerdem, E. Acar, "Soil moisture inversion via semiempirical and machine learning methods with full-polarization Radarsat-2 and polarimetric target decomposition data: A comparative study." *IEEE Access*, 8, 197896-197907, 2020.
- [16] W. Xu-dong, G. Jian-ming, J. Bai-jun et al., "Mixed-pixel classification of remote sensing images of cellular automata." *J Surv Mapp* 37(1):42-48, 2008.
- [17] G. Wu, X. Shao, Z. Guo, Q. Chen, W. Yuan, X. Shi, et al. "Automatic building segmentation of aerial imagery using multi-constraint fully convolutional networks", *Remote Sensing*, 10, p. 407, 2018.
- [18] J. Yuan, "Learning building extraction in aerial scenes with convolutional networks." *IEEE Transactions on Pattern Analysis Machine Intelligence*, 40, pp. 2793-2798, 2017.
- [19] Q. Chen, L. Wang, Y. Wu, G. Wu, Z. Guo, S. L. Waslander, "Aerial imagery for roof segmentation: A large-scale dataset towards automatic mapping of buildings." *ISPRS Journal of Photogrammetry and Remote Sensing*, 147, pp. 42-55, 2018.

- [20] <http://study.rsgis.whu.edu.cn/pages/download/>
- [21] O. Ronneberger, P. Fischer, T. Brox, "U-net: Convolutional networks for biomedical image segmentation." International conference on medical image computing and computer-assisted intervention, (pp. 234–241). Springer, 2015.
- [22] J. Long, E. Shelhamer, T. Darrell, "Fully Convolutional Networks for Semantic Segmentation", University of Berkeley, Proceedings of the IEEE, 2015.
- [23] N. Ketkar, "Stochastic gradient descent.", In Deep learning with Python (pp. 113-132). Apress, Berkeley, CA, 2017.
- [24] A. Radford, L. Metz, S. Chintala, "Unsupervised representation learning with deep convolutional generative adversarial networks.", arXiv preprint arXiv:1511.06434, 2015.
- [25] F. Chollet, "Keras: Deep learning library for theano and tensorflow", 2015, [online] Available: <https://github.com/fchollet/keras>.
- [26] Google Colab [Online] Access Link: <https://colab.research.google.com/>, on 21 November 2022.
- [27] G. Chhor, B. A. Cristian, B-L. Ianis, "Satellite image segmentation for building detection using U-Net." Web: <http://cs229.stanford.edu/proj2017/final-reports/5243715.pdf>, 2017.
- [28] İ. Ataş, "Human gender prediction based on deep transfer learning from panoramic dental radiograph images." Traitement du Signal, 39(5), 1585-1595, 2022. DOI:10.18280/ts.390515
- [29] A. H. Murphy, "The Finley Affair: A Signal Event in the History of Forecast Verification." Weather and Forecasting. 11 (1): 3, 1996.
- [30] Jaccard, Paul, "The Distribution of the Flora in the Alpine Zone.1". New Phytologist. 11 (2): 37–50, 1912. DOI:10.1111/j.1469-8137.1912.tb05611.x.
- [31] T. Sørensen, "A method of establishing groups of equal amplitude in plant sociology based on similarity of species and its application to analyses of the vegetation on Danish commons." Kongelige Danske Videnskabernes Selskab. 5 (4): 1–34, 1948.
- [32] L. R. Dice, "Measures of the Amount of Ecologic Association Between Species." Ecology. 26 (3): 297–302, 1945. DOI:10.2307/1932409.
- [33] F. Milletari, N. Navab, S. A. Ahmadi, "V-Net: Fully convolutional neural networks for volumetric medical image segmentation." In Proceedings of the 14th 3D Vision, Stanford, CA, USA, 25–28, pp. 565–571, 2016.
- [34] J. Zhang, et al. "Segmenting purple rapeseed leaves in the field from UAV RGB imagery using deep learning as an auxiliary means for nitrogen stress detection." Remote Sensing 12.9, 1403, 2020.
- [35] J. Ma, et al., "Building Extraction of Aerial Images by a Global and Multi-Scale Encoder Decoder Network." Remote Sens., 12, 2350, 2020.
- [36] J. Lin, W. Jing, H. Song, G. Chen, "ESFNet: Efficient Network for Building Extraction from High-Resolution Aerial Images." IEEE Access, 7, 54285–54294, 2019.
- [37] V. Iglovikov, A. Shvets, "Ternausnet: U-net with vgg11 encoder pre-trained on imagenet for image segmentation." arXiv preprint arXiv:1801.05746, 2018.
- [38] G. Chhor, C. B. Aramburu, I. Bougdal-Lambert, Satellite image segmentation for building detection using U-Net. Web: <http://cs229.stanford.edu/proj2017/final-reports/5243715>, 2017.
- [39] D. Patil, K. Patil, R. Nale, S. Chaudhari, "Semantic Segmentation of Satellite Images using Modified U-Net," IEEE 10. Regional Symposium (TENSYP), s.1-6, 2022.

BIOGRAPHIES

İSA ATAŞ has been an instructor with the Diyarbakır Vocational School of Technical Sciences, Dicle University. The research area is deep learning, computer vision, machine learning, and wireless networks.

Publication Ethics

The journal publishes original papers in the extensive field of Electrical-electronics and Computer engineering. To that end, it is essential that all who participate in producing the journal conduct themselves as authors, reviewers, editors, and publishers in accord with the highest level of professional ethics and standards. Plagiarism or self-plagiarism constitutes unethical scientific behavior and is never acceptable.

By submitting a manuscript to this journal, each author explicitly confirms that the manuscript meets the highest ethical standards for authors and coauthors

The undersigned hereby assign(s) to *Balkan Journal of Electrical & Computer Engineering* (BAJECE) copyright ownership in the above Paper, effective if and when the Paper is accepted for publication by BAJECE and to the extent transferable under applicable national law. This assignment gives BAJECE the right to register copyright to the Paper in its name as claimant and to publish the Paper in any print or electronic medium.

Authors, or their employers in the case of works made for hire, retain the following rights:

1. All proprietary rights other than copyright, including patent rights.
2. The right to make and distribute copies of the Paper for internal purposes.
3. The right to use the material for lecture or classroom purposes.
4. The right to prepare derivative publications based on the Paper, including books or book chapters, journal papers, and magazine articles, provided that publication of a derivative work occurs subsequent to the official date of publication by BAJECE.
5. The right to post an author-prepared version or an official version (preferred version) of the published paper on an internal or external server controlled exclusively by the author/employer, provided that (a) such posting is noncommercial in nature and the paper is made available to users without charge; (b) a copyright notice and full citation appear with the paper, and (c) a link to BAJECE's official online version of the abstract is provided using the DOI (Document Object Identifier) link.



ISSN: 2147- 284X
Year: January 2023
Volume: 11
Issue: 1

CONTENTS

Semiha Koşu, Serdar Özgür Ata; NOMA-Enabled Cooperative V2V Communications with Fixed-Gain AF Relaying,	1-12
Muhammet Fatih Aslan; Comparative Analysis of CNN Models and Bayesian Optimization-Based Machine Learning Algorithms in Leaf Type Classification,	13-24
Sadiq Iqbal, Jehad M. Hamamreh; Precoded Universal MIMO Superposition Transmission for Achieving Optimal Coverage and High Throughput in 6G and Beyond Networks,	25-34
Seda Üstün Ercan, Mohammed Sufyan Mohammed; IoT and XBee Based Central Car Parking Management System,	35-41
Ali Burak Öncül; LSTM-GRU Based Deep Learning Model with Word2Vec for Transcription Factors in Primates,	42-49
Sevilay Tüfenkçi, Barış Baykant Alagöz, Celaleddin Yeroğlu, Bilal Şenol; A Software Realization of Disturbance Rejection Optimal FOPID Controller Design Methodology by Using Soft Computing Techniques,	50-60
Tolga Ulutaş, Osman Avcı, Engin Can Akar, Barış Köksal, Yılmaz Kalkan; Simple Design and Implementation of Two-Way Communication System through UAV,	61-70
Nur Hüseyin Kaplan; Single Image Dehazing based on Additive Wavelet Transform,	71-77
Emrah Aydemir, Raghad Tohmas Esfandiyar Alalawi; Classification Of Hand Images by Person, Age and Gender with The Median Robust Extended Local Binary Model,	78-87
Hasan Hüseyin Coban; Analyzing the Societal Cost of Electric Roads Compared to Batteries and Oil for All Forms of Road Transport,	88-99
İsa Atas; Performance Evaluation of Jaccard-Dice Coefficient on Building Segmentation from High Resolution Satellite Images,	100-106

BALKAN JOURNAL OF ELECTRICAL & COMPUTER ENGINEERING

(An International Peer Reviewed, Indexed and Open Access Journal)

Contact

Batman University
Department of Electrical-Electronics Engineering
Bati Raman Campus Batman-Turkey

Web: <http://dergipark.gov.tr/bajece>
<http://www.bajece.com>
e-mail: bajece@hotmail.com

