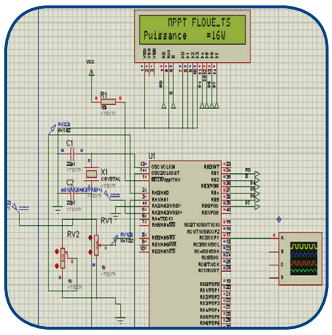
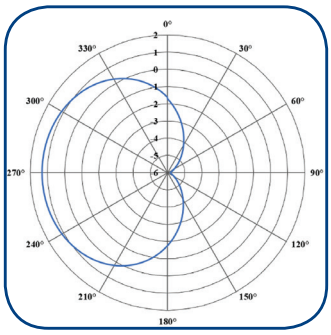
Volume 13, Number 10, 2022 ISSN 1847-6996

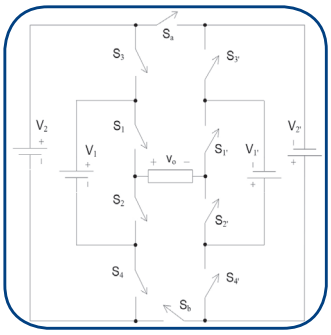
FERIT
FACULTY OF ELECTRICAL ENGINEERING, COMPUTER SCIENCE AND INFORMATION TECHNOLOGY **OSIJEK**

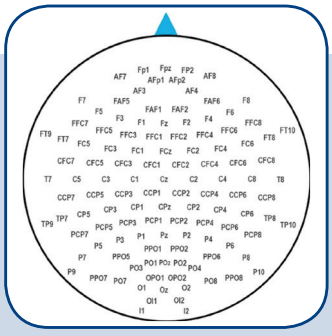
IJECES International Journal of Electrical and Computer Engineering Systems

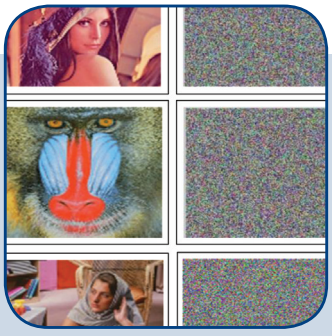
International Journal of Electrical and Computer Engineering Systems

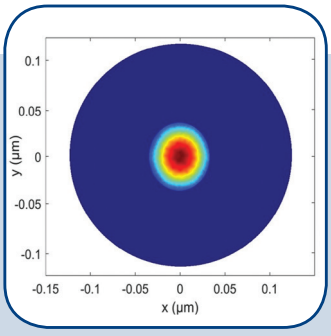












ISSN: 1847-6996

INTERNATIONAL JOURNAL OF ELECTRICAL AND COMPUTER ENGINEERING SYSTEMS

Published by Faculty of Electrical Engineering, Computer Science and Information Technology Osijek, Josip Juraj Strossmayer University of Osijek, Croatia

Osijek, Croatia | Volume 13, Number 10, 2022 | Pages 839 - 981

The International Journal of Electrical and Computer Engineering Systems is published with the financial support of the Ministry of Science and Education of the Republic of Croatia

CONTACT

International Journal of Electrical and Computer Engineering Systems (IJECES)

Faculty of Electrical Engineering, Computer Science and Information Technology Osijek,
Josip Juraj Strossmayer University of Osijek, Croatia
Kneza Trpimira 2b, 31000 Osijek, Croatia
Phone: +38531224600, Fax: +38531224605
e-mail: ijeces@ferit.hr

Subscription Information

The annual subscription rate is 50€ for individuals, 25€ for students and 150€ for libraries.
Giro account: 2390001 - 1100016777,
Croatian Postal Bank

Journal is referred in:

- Scopus
- Web of Science Core Collection (Emerging Sources Citation Index - ESCI)
- Google Scholar
- CiteFactor
- Genamics
- Hrčak
- Ulrichweb
- Reaxys
- Embase
- Engineering Village

Bibliographic Information

Commenced in 2010.
ISSN: 1847-6996
e-ISSN: 1847-7003
Published: quarterly
Circulation: 300

IJECES online

https://ijeces.ferit.hr

TABLE OF CONTENTS

A dual-band high gain complementary split-ring resonator (CSRR) loaded hexagonal bowtie antenna with enhanced bandwidth for Vehicle-to-Vehicle (V2V) communication applications

Original Scientific Paper

Rayan Hamza Alsisi
Faculty of Engineering, Department of Electrical Engineering,
Islamic University of Madinah, P.O. BOX 170 Madinah, 41411,
Saudi Arabia

Arshad Karimbu Vallappil
Faculty of Engineering, Department of Electrical Engineering,
Islamic University of Madinah, P.O. BOX 170 Madinah, 41411,
Saudi Arabia

Hafiz Abdul Wajid
Faculty of Engineering, Department of Electrical Engineering,
Islamic University of Madinah, P.O. BOX 170 Madinah, 41411,
Saudi Arabia

Abstract *– A highly reliable and efficient communication system is needed for a vehicle to navigate and drive to the destination without human control (known as an autonomous or self-driving vehicle). In this work, we consider various parameters for the antenna design, ensuring reliable communication amongst vehicles and infrastructure. Specifically, we consider the type of antenna, the method used, operating frequency, substrate type (with thickness and permittivity), size and shape, gain, and bandwidth. An optimal threshold value or range of these parameters is identified. Moreover, a complementary split-ring resonator (CSRR) metamaterial (MTM) based hexagonal bowtie antenna for a high gain V2V communication environment is presented. This antenna covers sub-6 GHz fifth generation (5G) bands (3.15-3.95 GHz) and Wi-Fi band 2.4GHz. Printing was done on a low-cost FR4 substrate for the radiating patch. Antenna Bandwidth is enhanced using a partial ground plane. The radiating layer is based on hexagonal patches printed on the double side of the substrate, and the CSSR structure is etched from patches to enrich antenna gain and bandwidth. More importantly, the proposed CSRR employed antenna provides gain and bandwidth of 1.6dBi / 6 dBi and 100MHz/ 8000MHz at 2.4GHz /3.5GHz, respectively. A highly known software, CST microwave studio, simulates the proposed antenna. Simulated and measured results make this arrangement a potential candidate for 5G high gain V2V communication.*

Keywords*: Autonomous vehicle, Optimal antenna parameters, Reliable and efficient communication systems, Metamaterial*

1. INTRODUCTION

Recent research findings conclude that: ***"nearly 1.25 million people die in road crashes each year, on average 3,287 deaths a day, and an additional 20-50 million are injured or disabled around the world"*** [1]. Though demand for automobile use in high numbers is unavoidable, the above numbers are not acceptable, especially if we know that most of these accidents happen due to human error. This poses a significant challenge for the whole world to redesign the existing transportation system irrespective of its serious problem.

The availability of fully Autonomous Vehicles (AVs) is one of the solutions not recently proposed instead in experimentation and evolutionary stages for a long time. The concept of AVs is becoming possible because of the advancements in multiple fields such as sensing, availability of state-of-the-art materials, ultra-fast computing facility, and Artificial Intelligence (AI). However, the underlying problem is not straightforward as it requires an upgraded intelligent transportation system where AVs are the main components [2,3]. Tons of efforts have been devoted in the past [4,5,6,7], which are remarkably commendable to enjoy the latest version

of AVs. However, it is surprising that no "fully autonomous vehicles" still exist [8]. Instead, the proposed contemporary solutions still reflect limitations and require improvement to enjoy an optimal version of AVs, specifically fast communication, to make reliable decisions fulfilling safety requirements [9,10].

In [11], a predictive active front steering control system for AVs is introduced where a trajectory is known for finite distance at each time step. Moreover, the front steering angle is determined for trajectory, especially for slippery roads at maximum entry speed. Furthermore, a map-free lane method was presented in [12], utilizing a 2D low-cost laser scanner for near future AVs to bridge the gap between forthcoming AVs and lane-keeping assistants. Moreover, integrated positioning for connected vehicles enhances positioning accuracy and availability, specifically for urban canyons studied in [13]. Furthermore, in [14], an advanced lane-keeping assistance system *(known as LKAS)* was proposed. In [15], a photovoltaic LIFI communication system offering both reliability and efficiency amongst AVs and between vehicles and road infrastructure (V2RI) was developed and termed as an intelligent transportation system keeping its significance compared with a traditional communication system using radio frequency (RF). Finally, it is worth mentioning [16] that an RF-based communication system covering a range of 20 kHz to 300 GHz was presented mainly based on the oscillation rate of an alternating electric current or voltage.

We now consider efforts devoted to wireless-based communication systems [17,18] specifically related to emerging intelligent transport systems (known as ITS). In this regard, a detailed survey is conducted in [17], highlighting what has already been contributed to wireless communication systems. Moreover, readers can enlighten further about methods a) used so far and b) currently under the research and validation phase.

It is further emphasized in [18] that ITS will fulfill the future needs of smart or intelligent communication systems. Consequently, humanity will witness safer highways or, in general, road travel. Here we name a few of the existing widely used and tested technologies such as Wi-Fi (IEEE 802.11xx), WiMAX (IEEE 802.16), 4G-LTE, wireless sensor, and ad hoc networks, and above all 5G. Concluding, the prime functionality of any communication system related to a transportation system is twofold. One is the critical communications between vehicles and infrastructure, addressing increased efficiency and reliable connectivity. Second is the passenger's experience ensuring utmost safety. In this regard, using antennas to communicate among vehicles and infrastructure has gained the attention of arrays and compact antennas [19-22]. a broadband high-gain [23-32], high gain and bandwidth [33], substrate material [34-40], high bandwidth [41-42].

This paper aims first to identify the optimal either threshold value or range for key performing parameters while designing an antenna and then design an antenna based on this for testing and validation. Achieving this, we ensure features such as efficiency, safety, and smooth navigation of AV. The structure of the paper is organized as follows. Section 2 is devoted to contributions made by many recently to identify optimal range or threshold values related to key performing parameters in antenna designing. Section 3 is devoted to designing and fabricating a novel complementary split-ring resonator (CSRR) metamaterial (MTM) based hexagonal bowtie antenna, followed by the conclusion in Section 4.

2. THRESHOLD/RANGE OF KEY PERFORMING PARAMETERS IN ANTENNA DESIGNING

In this section, we have primarily considered the following vital areas related to antenna design for V2V communication: array and compact antenna, broadband high-gain using high-end substrate material, high bandwidth, and high gain and bandwidth [19-22,33-42]. Commendably a lot is technically contributed, which is highlighted in the coming subsection for designing and implementing a novel antenna with the significance of optimal performance. Microstrip patch antennas (also known as planar antennas) are highly attractive as they are used in various applications and offer a range of features such as profile configuration, easy fabrication and manufacturing, cost-effectiveness, and lightweight.

2.1. HIGH GAIN ANTENNA USING WIDEBAND REFLECT ARRAY

In this section, we consider that another important class of antennas (*highly robust, reliable, and cost-effective with easy fabrication)* is the single layer fully planar reflect array antennas [19, 23-26]. Interestingly, these offer a much wider operating frequency range of [10-30 GHz] with 100% fractional bandwidth [19]. Moreover, the following salient features of the proposed single-layer array are highlighted below, which were thoroughly compared with related efforts made by many in [23-26]:

- *Wider operating frequency range of 10GHz to 45 GHz;*
- *Peak gain between range 24.2 to 32.83 dBi for the above-mentioned frequency range;*
- *bandwidth gain from 1 to 3 dB.*

2.2. HIGH GAIN ANTENNA USING HIGHER ORDER MODES

In [20], Govindarajulu et al., with the help of higher order modes, improved gain for an operating frequency of 28 GHz and an array operating at 5.9 GHz. The fabrication of the proposed aperture array for 5.9 GHz and 28 GHz operating frequencies ensured gains of 9.97 dBi and 12.3 dBi, respectively. Moreover, for

DSRC, band isolation of more than 55 dB amongst ports was seen, whereas, for 5G mm-wave, band isolation of 33 dB amongst ports was obtained. Following salient features of the single proposed layer array are highlighted below, which were thoroughly compared with related efforts made by many in [27-32]:

- Using higher order modes ensure high gain.
- Improvement in port-to-port isolation (*an inter-element spacing of 0.5λ*) for lower and upper band frequencies.
- Cost-effective and straightforward PCB-based fabrication because of its simple structure.

2.3. HIGH GAIN ANTENNA BY MODIFYING THE RADIATING PATCH

This section considers efforts devoted to high gain in connection with substrate material and modifying the radiating patch, which is of utmost importance while designing the antenna. We consider efforts invested in [34-39] based on different substrate materials with desired operating frequency band range of (5.88-5.92 GHz). Most of these articles use hexagonal antenna geometry with elliptical slots on the radiating patch, while V shape slots are used in [35]. In [39], gains of 4.37 dBi and 5.6 dBi were obtained without and with defective ground structure (DGS). However, amongst all efforts [34-39] highest gain of 6.02 dBi is ensured in [39]. Therefore, we mention below the salient features of this work:

- Offers the highest gain of 6.02 dBi;
- Antennas operate in the IEEE 802.11p defined vehicular communication band with 5.9 GHz resonant frequency and 200 MHz bandwidth;
- Geometry is less complex, enabling easy fabrication and cost-effectiveness.

2.4. HIGH GAIN AND HIGH BANDWIDTH ANTENNA

High gain and bandwidth antennas offer various attractive features, specifically in AVs applications. It is evident from [40-42] that higher order modes highlight enormous advantages, particularly when the antenna is designed using a single-layer simple patch. These antennas are found in diversified fields covering concealed weapons detection and enhanced vision for biomedical and avionics industries. In [42], optimal peak gain of 11.84 dBi with 17.6% bandwidth and gain of 8.5 dBi with 15.14% bandwidth for Ku and Ka bands is achieved so far, respectively.

2.5. COMPLEMENTARY METAMATERIAL RING RESONATOR (CSRR) ANTENNA

In this era, flat patch antennas with a suitable metal combination are highly desirable by the automotive industry covering high-frequency applications for V2V communication operating at a band of 5-6 GHz. We refer few of the efforts [43-45], where the reader can explore more from these texts. In [43] two complementary metamaterial resonators square patches were added to a rectangular patch antenna to optimize performance with a gain of 18.53 dB.

2.6. SUMMARY OF KEY FINDINGS

This section presents the summary of key findings from sections 2.1-2.5. it is evident from considered research works that efforts are devoted to improving gain using a high-end substrate, or by using an array antenna, and modifying the radiating patch. However, few of these research works focused only on improving bandwidth without enhancing considerable antenna gain. The major requirement for V2V communication antenna must provide both high gain and bandwidth by using low-cost substrate material for large-scale deployment related to vehicular applications [33]. Therefore, the following sections are devoted to incorporating above mentioned key requirements. Moreover, an optimal antenna design fulfilling characteristics such as better gain and bandwidth with a low-cost substrate has been simulated, fabricated, and validated at 3.5GHz.

3. DESIGN GUIDELINES AND STRUCTURE OF PROPOSED ANTENNA

The geometry of the proposed double-sided printed bowtie antenna is given in Fig.1 with the thickness and relative permittivity of the FR4 substrate as 0.8 and 4.3 mm respectively. Bowtie microstrip patch antennas became enticing candidates in modern communication systems due to their small size, which is smaller than that of a traditional rectangular radiating patch antenna at the same operating frequency. Moreover, the double-sided bowtie structures such as hexagonal and circular bowtie radiators help to improve the bandwidth from 10% to 50% [46,47]. So, this paper presents a modified double-sided bowtie design consisting of two hexagonal patches, one printed on the top layer and the other on the bottom layer (see Fig-1 a & b respectively). The double-sided patches have been etched with CSRR-MTM structure as shown in Fig. 1. The dimension and structure of CSRR are shown in Fig.2. Antenna is excited by a 50-Ω microstrip feed line. The step-by-step design of the proposed structure is explained below.

In Fig-3, we present a step-by-step design of the antenna. Initially, the design of the antenna starts with a hexagonal patch bowtie structure printed on the top layer and the bottom layer is the complete ground plane. The sides dimension of the hexagonal radiating patch has been developed from a circular patch because the shape of both patches is close relative to each other. The circular radiating patch antenna resonance frequency is determined by Equation (1) [48]. For TM11 mode, X_{mn} = 1.8411, a_e = effective radius of the

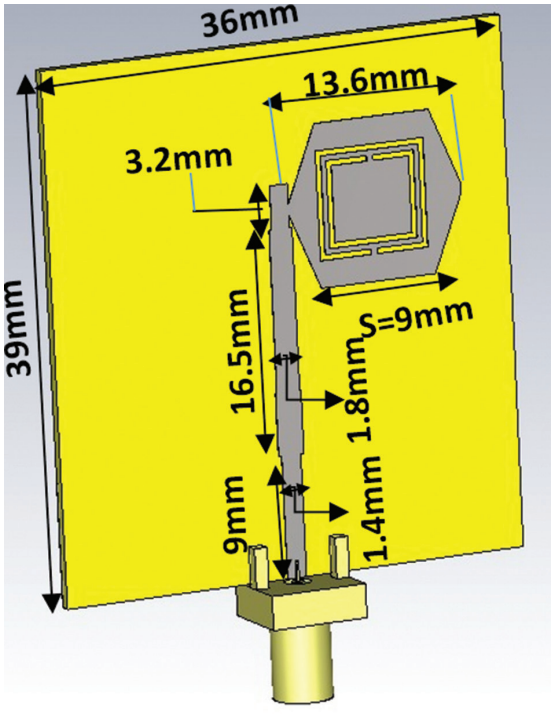
circular patch, and *c* is the free space velocity of light. In Step 1, the initial resonant frequency should be 5.8GHz. By substituting the value of f_r in Equation (1), the value of a_e =7.4mm. In Equation (2), *a* represents the circular patch's actual radius, whereas *h* indicates the thickness of the FR4 substrate. Equation (3) shows relations between circular and hexagonal radiating patch areas, whereas '*s*' represents hexagonal radiating patch side length. Equation (1)-(3) has been used to calculate the '*s*'.

$$f_r = \frac{X_{mn}}{2\pi a_e \sqrt{\varepsilon_r}} c \quad (1)$$

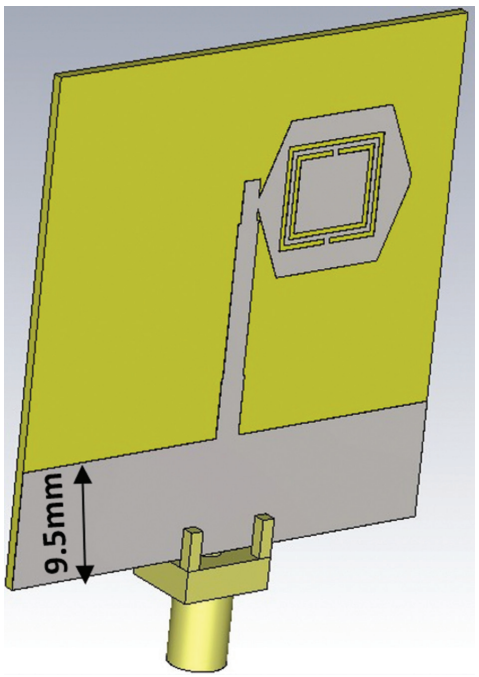
$$a_e = a\left\{1 - \frac{2h}{\pi a \varepsilon_r}\left(ln\frac{\pi a}{2h} + 1.7726\right)\right\}^{0.5} \quad (2)$$

$$\pi a_e^2 = \frac{3\sqrt{3}}{2} s^2 \quad (3)$$

Likewise, it has been shown in Fig.4 that the antenna resonates 5.8 GHz in step 1 with a bandwidth of 350 MHz by simulating the design using CST Microwave Studio. In Step 2, a CSRR structure was etched from the hexagonal patches. This helps to shift the resonating frequency from 5.5 GHz to 4.3 GHz and generated a new resonant frequency at 2GHz. So, the antenna shows a dual-band behaviour. In Step 3, a bandwidth improvement was achieved by moving one of the hexagonal patches to the bottom layer with a partial ground plane and the other one to the top layer. An improved bandwidth of 100MHz / 850 MHz has been achieved at 2.4GHz /3.5GHz, respectively. The final design of the proposed antenna covers a dual-band frequency range between 2.35 GHz- 2.45 GHz and 3.2 GHz- 4.05GHz. The first resonant frequency of the proposed antenna covers the Wi-Fi band, and the second frequency covers the sub-6GHz 5G frequency.



(a)



(b)

Fig. 1. Layout of Proposed Antenna Design (a) Top Layer (b) Bottom Layer

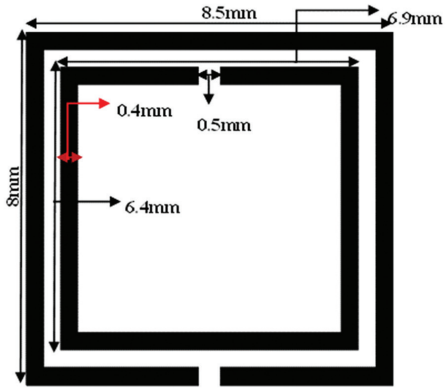


Fig. 2. Layout of Complimentary Split Ring Resonator 11

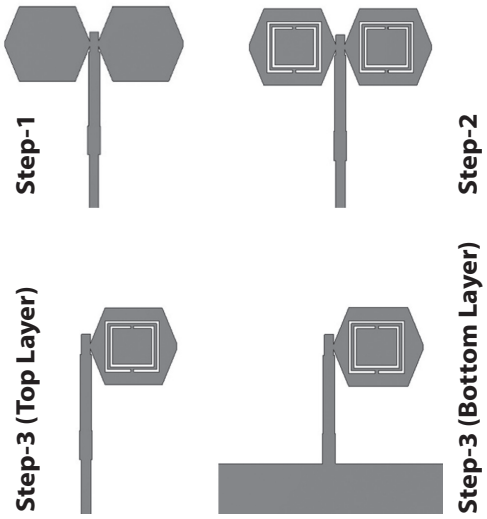


Fig. 3. Step-by-Step Design of Proposed antenna

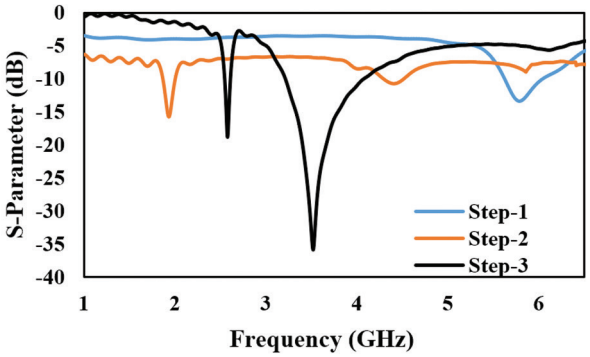
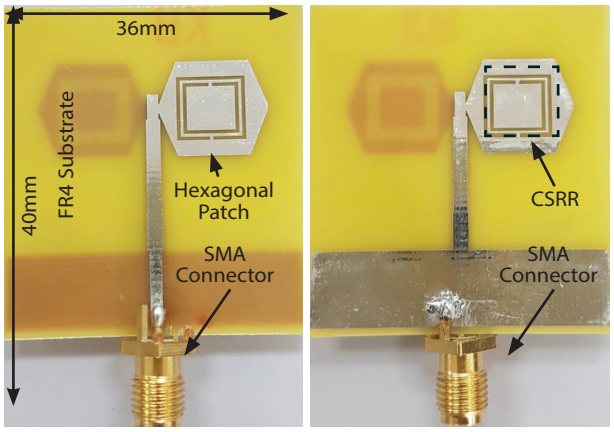


Fig.4: S- Parameter response of step-by-step antenna design

4. RESULT AND DISCUSSION

The fabricated prototype of the proposed CSRR MTM-based double-sided hexagonal bowtie antenna is shown in Fig.5 (a-b). The simulated and measured results of the return loss (S_{11}) for the proposed bowtie antenna are shown in Fig. 6.



(a) (b)

Fig.5. Fabricated prototype of Proposed Antenna Design (a) Top Layer (b) Bottom Layer

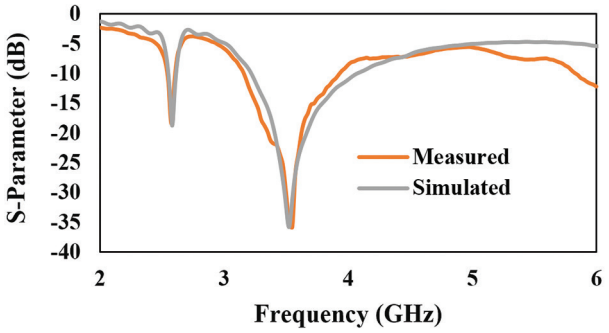
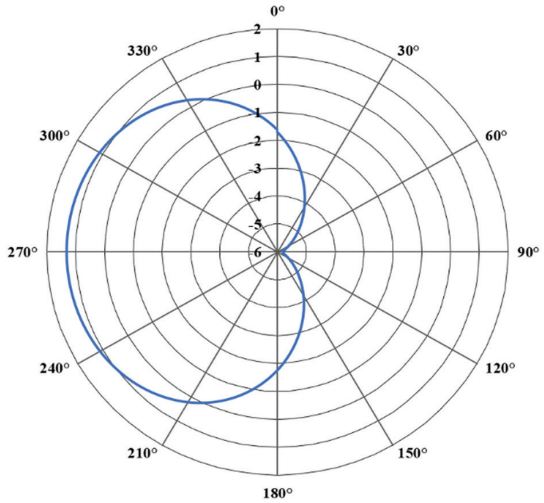


Fig.6. S-Parameter Response of Proposed Antenna

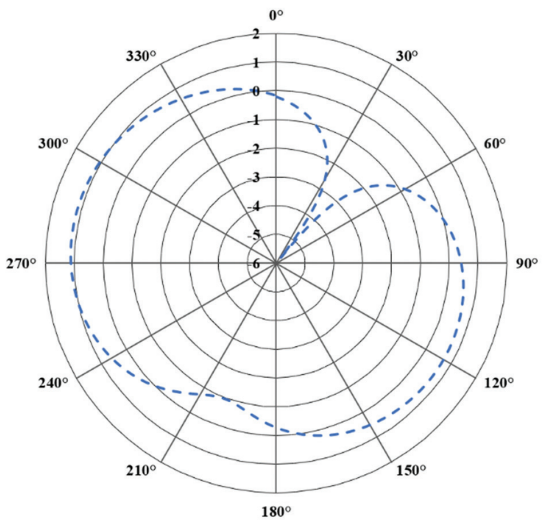
All simulated results were computed using the CST Microwave Studio which works on the FDTD technique. It could be observed from Fig. 6 that the proposed antenna resonates at 2.4GHz and 3.5GHz with a return loss of -20dB and -40dB, respectively.

Furthermore, it is observed that the measured results of the proposed antenna cover the dual-band frequency range of 2.38GHz- 2.48GHz and 3.15GHz -3.95GHz. The S_{11} values suggest good impedance matching at 50 Ω and reasonable bandwidth. The S-parameter response shows an excellent agreement between simulated and measured results.

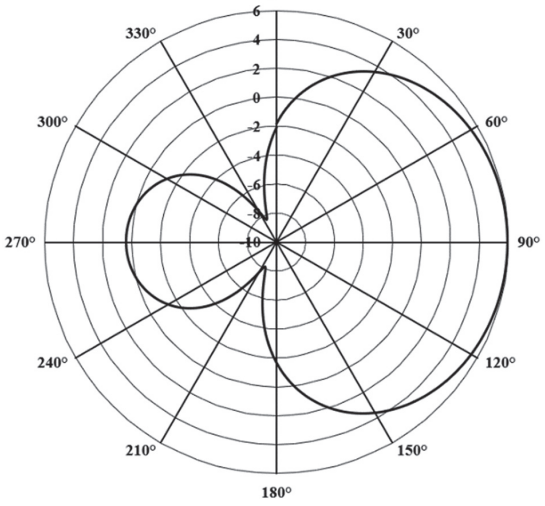
Fig.7 (a-d) shows the proposed antenna's simulated and measured radiation patterns at both the E-plane and H-plane, respectively. It can be observed from **Fig.7 (a-b)** that the proposed bowties antenna is radiating in the direction of -90° having a gain of 1.6dBi at 2.4GHz. Similarly, it has been illustrated in the **Fig.7 (c-d)** that the antenna is radiating in the direction of 90° with a gain of 6dBi at 3.5GHz. At higher frequency, the antenna gain is more as compared to lower frequency and the direction is opposite to each other. This shows that the proposed antenna is an excellent candidate for V2V communication due to its high gain at 3.5GHz and covering two different directions at the same time in two different frequencies. The high gain at 3.5GHz makes the proposed antenna design more suitable to work in the 5G wireless communication systems.



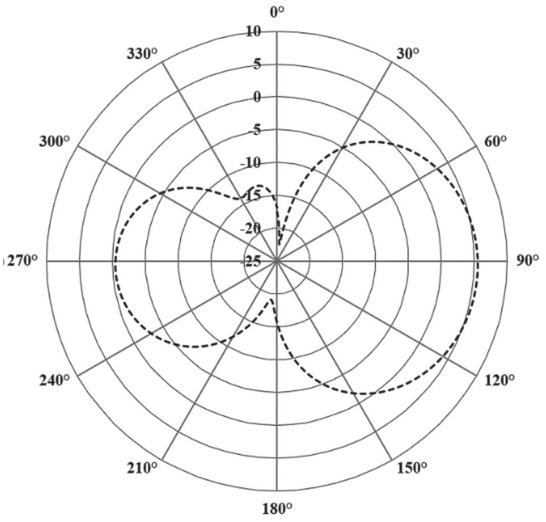
(a)



(b)



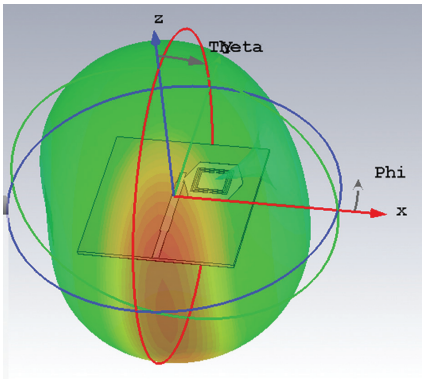
(c)



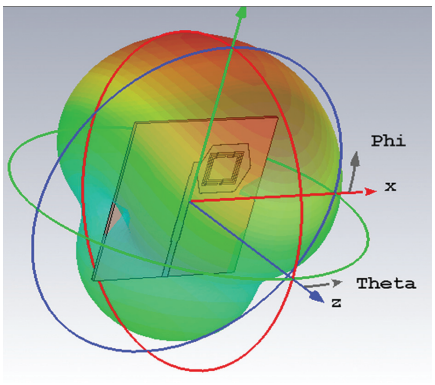
(d)

Fig.7. Two-Dimensional Radiation Pattern of Proposed Antenna (a) E-Plane (b) H-Plane at 2.4GHz (c) E-Plane (d) H-Plane at 3.5GHz

The 3D radiation pattern of the proposed antenna at 2.4GHz and 3.5GHz is shown in **Fig.8 (a-b)**. It can be seen from Fig.8 that the proposed antenna shows that maximum radiation at 2.4GHz and 3.5GHz are opposite to each other. This shows that the proposed antenna was a promising candidate for V2V communication.



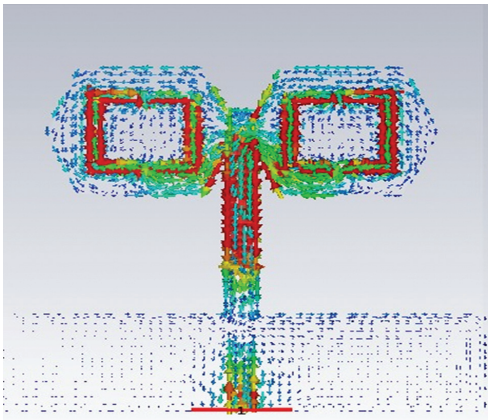
(a)



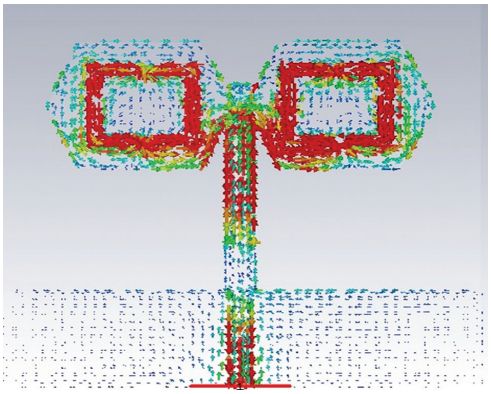
(b)

Fig.8. 3D Radiation Pattern of Proposed Antenna (a) 2.4GHz (b) 3.5GHz

After measuring the 3D radiation patterns, the surface current distribution of the designed and proposed antenna is simulated to understand the antenna behavior. **Fig.9 (a-b)** shows the antenna surface current response at 2.4GHz and 3.5GHz, respectively. It can be observed from the results in **Fig.9 (a-b)** that the surface current distribution is high around the CSRR structure and low at the corners. **Fig.10** shows the simulated radiating efficiency of the proposed antenna. It can be realized from Fig.10 that the proposed antenna shows a radiating efficiency of 90% at 3.5GHz.



(a)



(b)

Fig.9. Proposed antenna surface current distribution at (a) 2.4GHz and (b) 3.5GHz, respectively

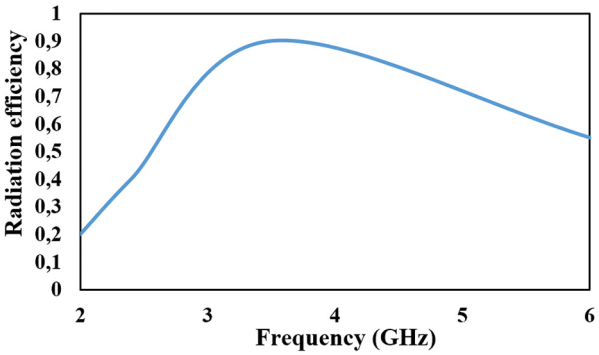


Fig.10. Radiating Efficiency of Proposed Antenna

Table 1. Comparison of Proposed antenna design with the existing designs [34,44, 49-51]

Ref	Centre Frequency (GHz)	Single / Dual-band	10dB -Bandwidth (%)	Gain (dBi)	Size (mm × mm)
[34]	5.9	Single	3	5.6	20.2 × 24.1
[44]	2.45, 3.6	Dual	10, 8	3	35 × 10
[49]	3.5	Single	12.57	7	37.4 × 37.4
[50]	3.5, 5	Dual	14, 28	2.5, 4.1	32 × 20
[51]	5.8	Single	8.27	6.15	200 × 200
This Work	2.4, 3.5	Dual	4, 23	1.6, 6	36 × 39

A comparison of the proposed antenna with the previously published work on antennas for vehicular application is summarized in Table 1. The researchers are attempting to either enhance the bandwidth or improve the gain at 3.5 GHz/ 5.9GHz in most of the designs described in Table 1. The proposed work shows the bandwidth and gain improvement of the antenna together at 3.5GHz, which is considered the main requirement for the antenna design of V2V Communication.

5. CONCLUSION

Smart cities are becoming a necessity for the whole world due to their benefits, for which one of the main components is intelligent transportation systems and AVs. Keeping this in mind, famous car makers, Google, Apple, and Uber, are investing huge amounts of money. As a result, the autonomous and semi-autonomous vehicle market is expected to reach $500 billion by 2035 [15]. Furthermore, AVs are considered the best solution for the increased number of accidents happening due to human errors. Therefore, a noticeable improvement to autonomous vehicles and transportation systems is developed by designing a smart communication process between the vehicles and the surroundings to guarantee safety. One of the key components in this regard is the antenna which plays a vital role in reliable communication. Considering this, a CSRR MTM-based hexagonal bowtie antenna is presented to enjoy better gain and bandwidth, specifically for V2V applications. The proposed antenna covers sub-6 GHz fifth generation (5G) bands (3.15-3.95 GHz) and Wi-Fi band 2.4GHz with radiating patch printed on a low-cost FR4 substrate. Furthermore, the proposed antenna achieved a gain and bandwidth of 1.6dBi / 6 dBi and 100MHz/ 800MHz at 2.4GHz /3.5GHz. Finally, simulated and measured results make this arrangement a potential candidate for 5G high gain Vehicle-to-Vehicle (V2V) communication.

6. ACKNOWLEDGMENTS

The authors would like to acknowledge the funding from the Deanship of Scientific Research, Islamic University of Madinah, Madinah, Saudi Arabia, under Research Grant No. 650.

Funding: This research was funded by "DEANSHIP OF SCIENTIFIC RESEARCH, ISLAMIC UNIVERSITY OF MADINAH, MADINAH, SAUDI ARABIA, Grant Number 650".

7. REFERENCES

[1] Road Safety Facts-Association for Safe International Road Travel, http://www.asirt.org (accessed: 2019)

[2] B. Bechadergue, "Mesure de distance et transmission de données inter-véhicules par phares à LED", Université Paris-Saclay, 2017.

[3] What is an Autonomous Car? https://www.techopedia.com/definition/30056/autonomous-vehicle (accessed: 2019)

[4] J. B.Greenblatt, S. Shaheen. "Automated vehicles, on-demand mobility, and environmental impacts", Current Sustainable/Renewable Energy Reports, Vol. 2, No. 3, 2015, pp. 74-81.

[5] S. D. Pendleton et al. "Perception, planning, control, and coordination for autonomous vehicles", Machines, Vol. 5, No. 1, 2017, pp. 1-54.

[6] S. E. Shladover, D. Su, X. Y. Lu, "Impacts of cooperative adaptive cruise control on freeway traffic flow", Transportation Research Record, Vol. 2324, No. 1, 2012, pp. 63-70.

[7] A. Faisal et al. "Understanding autonomous vehicles: A systematic literature review on capability, impact, planning and policy", Journal of Transport and Land Use, Vol. 12, No. 1, 2019, pp. 45-72.

[8] K. Upamanyu, I. P. Ramaswamy, "The Effects of Image Augmentation on Efficiency of a Convolution-

al Neural Network of a Self-Driving Car", Journal of Student Research, Vol. 10, No. 2, 2021, pp. 1-15.

[9] Koopman P, Wagner M, "Autonomous vehicle safety: An interdisciplinary challenge", IEEE Intelligent Transportation Systems Magazine, Vol. 9, No. 1, 2017, pp.90-96.

[10] J. Wang et al. "Safety of Autonomous Vehicles", Journal of Advanced Transportation, Vol. 2020, 2020, pp. 1-13.

[11] P. Falcone, et al. "Predictive Active Steering Control for Autonomous Vehicle Systems", IEEE Transactions on Control Systems Technology, Vol. 15, No. 3, 2007, pp. 566-580.

[12] S. Zhiwei et al. "Map free lane following based on low-cost laser scanner for near future autonomous service vehicle", Proceedings of the IEEE Intelligent Vehicles Symposium (IV), Seoul, Korea, 28 June 2015, pp. 706-711.

[13] A. Mahmoud, A. Noureldin, H. S. Hassanein, "Integrated Positioning for Connected Vehicles", IEEE Transactions on Intelligent Transportation Systems, Vol. 21, No. 1, 2019, pp. 397-409.

[14] Y. Bian et al. "An Advanced Lane-Keeping Assistance System With Switchable Assistance Modes", IEEE Transactions on Intelligent Transportation Systems, Vol. 21, No. 1, 2019, pp. 385-396.

[15] D. George, S. Vaidyanathan, A. S. Rajput, K. Deepa, "LiFi for vehicle to vehicle communication–a review", Procedia Computer Science, Vol. 165, 2019, pp.25-31.

[16] J. A. Fleming, "The Principles of Electric Wave Telegraphy and Telephony, London", Longmans, Green & Co., 1916.

[17] M. Nakagawa, L. B. Michael, "Wireless Communication Technology in Intelligent Transport Systems", Wireless Communications, Springer-Boston, 1997, pp. 491-508.

[18] C. Briso-Rodríguez et al. "Wireless Communications in Transportation Systems", Wireless Communications and Mobile Computing, Vo. 2017, 2017, pp.1-2.

[19] L. Zhang et al. "A Single-Layer 10-30 GHz Reflectarray Antenna for the Internet of Vehicles", IEEE Transactions on Vehicular Technology, Vol. 71, No. 2, 2021, pp.1480-1490.

[20] S. R. Govindarajulu et al. "Low Profile Dual-Band Shared Aperture Array for Vehicle-to-Vehicle Communication", IEEE Access, Vol. 9, 2021, pp. 147082-147090.

[21] L. Qin et al. "Millimeter-wave slotted waveguide array with unequal beam widths and low side lobe levels for vehicle radars and communications", IEEE Transactions on Vehicular Technology, Vol. 67, No. 11, 2018, pp. 10574-10582.

[22] M. Mosalanejad et al. "Wideband compact combline antenna array for 79 GHz automotive radar applications", IEEE Antennas and Wireless Propagation Letters, Vol. 17, No. 9, 2018, pp. 1580-1583.

[23] Z. W. Miao, Z. C. Hao, "A Wideband Reflectarray Antenna Using Substrate Integrated Coaxial True-Time Delay Lines for QLink-Pan Applications", IEEE Antennas and Wireless Propagation Letters, Vol. 16, 2017, pp. 2582-2585.

[24] C. Han, Y. Zhang, Q. Yang, "A Broadband Reflect array Antenna Using Triple Gapped Rings with Attached Phase-Delay Lines", IEEE Transactions on Antennas and Propagation, Vol. 65, No. 5, 2017, pp. 2713-2717.

[25] H. Luyen, Z. Yang, M. Gao, J. H. Booske, N. Behdad, "A Wideband, Single-Layer Reflectarray Exploiting a Polarization Rotating Unit Cell", IEEE Transactions on Antennas and Propagation, Vol. 67, No. 2, 2019, pp. 872-883.

[26] W. Li et al. "An Ultra-Wide-Band Tightly Coupled Dipole Reflectarray Antenna", IEEE Transactions on Antennas and Propagation, Vol. 66, No. 2, 2018, pp. 533-540.

[27] J. D. Zhang, W. Wu, D. G. Fang, "Dual-band and dual-circularly polarized shared-aperture array antennas with single-layer substrate", IEEE Transactions on Antennas and Propagation, Vol. 64, No. 1, 2016, pp. 109-116.

[28] C. X. Mao, S. Gao, Y. Wang, Q. X. Chu, X. X. Yang, "Dual-band circularly polarized shared-aperture array for C-/X-band satellite communications", IEEE Transactions on Antennas and Propagation, Vol. 65, No. 10, 2017, pp. 5171-5178.

[29] Z. Shi-Gang, Y. Jiang-Jun, C. Tan-Huat, "Design of L/X-band shared aperture antenna array for SAR application", Microwave and Optical Technology Letters, Vol. 57, No. 9, 2015, pp. 2197-2204.

[30] M. Ikram et al., "Integrated frequency reconfigurable slot antenna and connected slot antenna array for 4G and 5G mobile handsets", IEEE Transactions on Antennas and Propagation, Vol. 67, No. 12, 2019, pp. 7225-7233.

[31] M. Ikram, M. S. Sharawi, A. Shamim, A. Sebak, "A multiband dual standard MIMO antenna system based on monopoles (4G) and connected slots (5G) for future smart phones", Microwave and Optical Technology Letters, Vol. 60, No. 6, 2018, pp. 1468-1476.

[32] B. J. Xiang, S. Y. Zheng, H. Wong, Y. M. Pan, K. X. Wang, M. H. Xia, "A flexible dual-band antenna with large frequency ratio and different radiation properties over the two bands", IEEE Transactions on Antennas and Propagation, Vol. 66, No. 2, 2018, pp. 657-667.

[33] M. Ignatenko et al. "Wide-Band High-Frequency Antennas for Military Vehicles: Design and testing low-profile half-loop, inverted-L, and umbrella NVIS antennas", IEEE Antennas and Propagation Magazine, Vol. 58, No. 6, 2016, pp. 64-74.

[34] V. R. Kanni, R. Brinda, "Design of High Gain Microstrip Antenna for Vehicle to Vehicle Communication Using Genetic Algorithm", Progress In Electromagnetics Research M, Vol. 81, 2019, pp. 167-179.

[35] H. Hao, J. Li, D. Huang, W. Luo, "Design of hexagon microstrip antenna for vehicle-to-vehicle communication", The Journal of China Universities of Posts and Telecommunications, Vol. 23, No. 4, 2016, pp. 69-76.

[36] T. Mondal, S. Maity, R. Ghatak, "Compact circularly polarized wide-beam width fern-fractal-shaped microstrip antenna for vehicular communication", IEEE Transactions on Vehicular Technology, Vol. 67, No. 6, 2018, pp. 5126-5134.

[37] A. Liu, Y. Lu, L. Huang, "Low-profile patch antennas with enhanced horizontal Omni-directional gain for DSRC applications", IET Microwaves, Antennas & Propagation, Vol. 12, No. 2, 2018, pp. 246-253.

[38] G. Z. Rafi et al., "Low-profile integrated microstrip antenna for GPS-DSRC application", IEEE Antennas and Wireless Propagation Letters, Vol. 8, 2009, pp. 44-48.

[39] H. Singhal et al. "High gain hexagonal patch antenna for v2v communication", Proceedings of the 7th International Conference on Signal Processing and Integrated Networks, Noida, India, 27-28 February 2020.

[40] Z. X. Xia, K. W. Leung, P. Gu, R. Chen, "3-D-Printed Wideband High-Efficiency Dual-Frequency Antenna for Vehicular Communications", IEEE Transactions on Vehicular Technology, Vol. 71, No. 4, 2022, pp. 3457-3469.

[41] Z. P. Zhong et al. "A Compact Dual-Band Circularly Polarized Antenna with Wide Axial-Ratio Beamwidth for Vehicle GPS Satellite Navigation Application", IEEE Transactions on Vehicular Technology, Vol. 68, No. 9, 2019, pp. 8683-8692.

[42] Q. U. Khan, M. B. Ihsan, D. Fazal, F. M. Malik, S. A. Sheikh, M. Salman, "Higher Order Modes: A Solution for High Gain, Wide Band Patch Antennas for Different Vehicular Applications", IEEE Transactions on Vehicular Technology, Vol. 66, No. 5, 2017, pp. 3548-3554.

[43] M. Berka et. al. "Smart metamaterial rectangular antenna design for 5g applications dedicated to autonomous vehicles", Proceedings of the 6th International Conference on Advances in Mechanical Engineering, Istanbul, Turkey, 20-22 October 2021, pp. 1-6.

[44] E. M. Malathy et al. "Miniaturized dual-band metamaterial-loaded antenna for heterogeneous vehicular communication networks", IETE Journal of Research, Vol. 2021, 2021, pp.1-10.

[45] A. S. M. Alqadami et al. "Multi-band antenna array based on double negative metamaterial for multi automotive applications", Progress In Electromagnetics Research, Vol. 159, 2017, pp. 27-37.

[46] K. Siakavara, "Methods to design microstrip antennas for modern applications", Microstrip Antennas, Intech, 2011, pp. 173-236.

[47] A. Dastranj, "Modified end-fire bowtie antenna fed by microstrip line for wideband communication systems", Journal of Electromagnetic Waves and Applications, Vol. 32, No. 13, 2018, pp.1629-1643.

[48] N. Kushwaha, R. Kumar, "Design of the slotted ground hexagonal microstrip patch antenna and gain improvement with FSS screen", Progress in Electromagnetics Research B, Vol. 51, 2013, pp. 177-199.

[49] H. Wang, G. Yang, "Design of 4 × 4 microstrip Quasi-Yagi beam-steering antenna array operation at 3.5 GHz for future 5G vehicle applications", Proceedings of the International Workshop on Antenna Technology: Small Antennas, Innovative Structures, and Applications, 2017, pp. 331-334.

[50] Y. Zhang et. al, "A compact dual-band antenna for 5G application", Proceedings of the Cross-Strait Quad-Regional Radio Science and Wireless Technology Conference, Taiyuan, China, 18-21 July 2019, pp. 1-2.

[51] D. Rongas et al. "An integrated shark-fin reconfigurable antenna for V2X communications", Progress in Electromagnetics Research C, Vol. 100, 2020, pp. 1-16.

Analog Feedback Communication System with Receive Diversity and MRC

Original Scientific Paper

Richa Tengshe
Department of Electronics and Communication Engineering,
CMR Institute of Technology, Bengaluru, Karnataka, India
richatengshe@gmail.com

Navin Kumar
Amrita School of Engineering Bengaluru,
Amrita Vishwa Vidyapeetham India
navinkumar@ieee.org

Abstract *– This study investigates the analog feedback communication system (AFCS) architecture considering Rayleigh fading channel model and receive diversity maximum ratio combining (MRC). This architecture employs a power-efficient transmitting unit, an estimator at the receiver side and iterative algorithm that minimizes the mean square error (MSE). Using the feedback channel, the estimated received sample is fed back to the transmitter. The performance of AFCS with the Rayleigh fading channel model is evaluated using MSE optimization. The investigation revealed that when compared to a single input single output AFCS system, the diversity-enabled AFCS system achieves negligible MSE in fewer iterations. MSE of order 10^{-3} is achieved by 6 receive antennas with MRC in only 4 iterations at 0dB channel signal-to-noise ratio (SNR), compared to a single input single output AFCS system that requires more than 10 iterations to achieve the same order of MSE.*

Keywords*: Shannon Capacity, Analog Feedback Communication System, Receive Diversity, Maximum Ratio Combining, Optimum Performance Theoretically attainable*

1. INTRODUCTION

A prevailing trend in wireless communication is increasing data rates, reliable and long-range transmission while reducing transmitter sizes, energy usage, and the costs [1]. This task is internally incoherent since using low-power transmitters to save energy consumption results in decreased quality, range, and transmission rate. Many foundational works in information theory and communication theory depict the highest possible performance limits for the transmission [2], [3]. One such milestone work by Shannon showed that transmission of information at the highest channel capacity is possible in the presence of noise [4]. However, no strict analytical guidelines were provided for the actual implementation of such systems which work on the performance bounds. With the invention of turbo codes and sub- sequent developments [5, 6], some information theory promises have more or less become a reality. However, the most popular source and channel coding schemes do not offer performance even close to ideal when the block lengths are short. With large block lengths, performance is almost ideal but the cost is paid in terms of unconstrained power, delay and complexity.

Another problem with the separation theorem based digital systems is that whenever we want to change the code rate or the distortion target, we must completely redesign the digital system. Furthermore, digital systems are subject to the cliff effect [7]. The system's performance suffers greatly if the channel distortions are less than the designed values, while it improves only slightly if the channel condition improves.

Analog communications, which rely on the transmission of discrete-time continuous-amplitude sources can be viewed as a viable alternative to digital systems. Analog communication is well known to perform best under certain conditions, such as direct transmission of Gaussian samples over AWGN channels with absolutely no coding required [8]. P. Elias demonstrated [9] in the middle of the 1950s that if the information source is statistically matched to the channel, the analog communication systems can convey signals at the information limits without complex coding if a feedback channel is available.

Feedback does not improve the capacity of a memory less channel [3], but is definitely known to reduce error exponent and the coding complexity needed for achieving the performance limits [10]. This finding

sparked intense curiosity among researchers and led to many significant works like [11]. The outcomes of these studies unequivocally showed that it is possible to design a perfect Analog Feedback Communication System (AFCS) whose performance reaches the limits of information and the complexity of AFCS systems is very less compared to the performance achieving digital systems. The peculiarity of these studies is that it directly optimizes the transmitter and receiver for a particular channel model, unlike current digital communication systems (DCS) where the source coder and channel coder optimization is done separately. However, towards the beginning of the 1970s, research in this area was reduced to almost nil and DCS became popular because of several advantages including ease of implementation, cost-effectiveness, and design. However, the research interest in AFCS has increased again, particularly for some applications like power efficient wireless sensor networks (WSN) [12, 13], small satellites [14, 15], radio frequency identification (RFID) modules, etc. AFCS architecture with a transmission and reception algorithm, an adaptive modulator, an estimator and a feedback channel are proposed and discussed in depth in works [16] [17]. Here MSE is emphasized to be the performance criteria and the relation of the optimality of the algorithm to the number of iterations is derived and presented. The research [18] and [19], demonstrated the possibilities and benefits of using feedback communication systems (FCS) as an estimating system. They discussed FCS optimization that eliminates the issues that led to the discontinuation of AFCS research and the solutions to those issues which allow for the design of the most efficient systems working at Shannon's bounds. Based on AFCS optimization, power-bandwidth efficiency and bit-rate expressions are discussed in [20]. These studies re-emphasized that the transmission over the "threshold" number of cycles n offers the "perfect" P-B trade-off and full utilization of the system's resources. Apart from the works mentioned so far, authors have discussed many resource-constrained use cases where AFCS might prove better compared to DCS [13], [14].

In All the research works carried out in the field of AFCS like [16] so far assumed the channel model to be additive white Gaussian (AWG). The real world wireless channel can be modelled better with Rayleigh or Rician fading channel model. The fading channel model and the techniques like diversity and multiple input multiple outputs (MIMO) to combat fading channel noise still remain unearthed and open for research. Hence, in this work we present the architecture of AFCS for wireless flat fading channels as a follow-up AFCS works. This is very important to maintain the research in this area and to come up with some practically implementable solution. The main contributions of this work are:

1. AFCS transmission and reception algorithm and mathematical details, considering wireless fading channel model and receive diversity.
2. Analysis of the effect of diversity and Maximum Ratio Combining (MRC) technique on the MSE performance for the number of receive antennas $N = 1, 2, 4, 6$.
3. Comparison of the wireless AFCS system's spectral efficiency for the different number of antennas ($N = 2, 4, 6$) with DCS with Phase Shift Keying (PSK) modulation scheme.

Following is an outline of the paper's content. The AFCS system architecture is explained in section 2. Section 3 discusses the detailed mathematical background, AWGN channel model, optimization algorithm, and mathematical analysis including fading channel model and diversity combining technique. In section 4, performance characteristics of AFCS are analyzed. Results and discussion is presented in section 5. Finally, section 6 provides the conclusion.

2. SYSTEM ARCHITECTURE

The AFCS schematic is shown in Fig. 1. It shows a system for the transmission of discrete time continuous amplitude memory-less Gaussian source over a wireless channel. The system is assumed to have a forward channel (ChF) and a feedback channel TxT2-ChR-RxR2. These channels connect the analog transmitting unit to the base station (receiver). We assume that the forward channel is a flat Rayleigh fading channel and the feedback channel (ChR) is a high-quality AWGN channel. A high quality noiseless feedback channel can be ensured by increasing the power transmitted from the receiver on the feedback channel. We presume the absence of memory for both channels. The AWG noise ζ_t in the forward channel is assumed to have a variance σ^2_{ζ}. The variance of feedback errors at the subtraction unit in the transmitting unit is considered to be σ^2_{v}. Without sacrificing generality, we assume that the source generates samples with a zero mean and variance σ^2_{0}.

In the overall system, the transmitting unit consists of a source, sample and hold circuit (Sampler), sub- traction unit, modulator (M1), and radio frequency (RF) transmitter module in the forward path. Along with this, it has a radio front-end receiving module and a demodulator (included in front-end receiver) as shown in Fig. 1 as a part of the Base Station (BS) unit. Additionally, BS unit has an estimator module (EM) in the forward path and a modulator and radio frequency transmitting system (TxT2) in the feedback path.

During the time interval T, each sample of the input signal is maintained at the subtraction unit's input and transmitted in n cycles (iterations), independent of the previous samples. BS unit analyses the signal received from the transmitting unit in each k^{th} cycle ($k = 1, 2, \ldots, n$) and computes the sam ple's intermediate estimate in the estimator module (EM) and saves this estimate till the next cycle. The control signal that is conveyed to the transmit- ting unit via the feedback channel is also computed by EM. The maximum number of cycles or it-

erations in which one sample of input can be sent is $n = T/(\Delta t_0) = F_0/F$ where $T = 1/2F$ is the sampling period and $\Delta t_0 = 1/(2F_0)$ is the duration of one cycle of transmission. The minimum bandwidth re- quired of the forward channel is $F_0 = 1/(2\Delta t_0)$ and the bandwidth of the feedback channel is assumed to be greater than F_0.

For each k^{th} cycle of transmission $(k = 1; 2; \ldots; n)$;

$$e_k = x_k - \hat{B}_k = x_k - \hat{x}_{k-1} + v_k, \tag{1}$$

where v_k is feedback error (AWGN) with variance σ^2_v. The transmitting unit modulator uses a double side band suppressed carrier pulse amplitude modulator (PAM). Every cycle, the values provided by the solution of the optimization problem are used to set the adaptive modulation depth of the modulator M_k. The transmitter has a nonlinear transfer function. The saturation appears if the signal $M_k e_k$ surpasses the carrier amplitude's saturation level. The transmitted signal is given by:

$$y_k = A_0 \begin{cases} M_k e_k, & \text{if } M_k|e_k| \le 1, \\ sign(e_k) & \text{if } M_k|e_k| > 1, \end{cases} \tag{2}$$

and the received signal is given by:

$$\tilde{y}_k^r = h y_k + \zeta_k, \tag{3}$$

where h is the fading coefficient and k is the Gaussian noise in the forward channel. $\tilde{y}_k^r$ is demodulated and the demodulated signal is given by:

$$\tilde{y}_k = A_0 h \begin{cases} M_k e_k + \zeta_k, & \text{if } M_k|e_k| \le 1, \\ sign(e_k) + \zeta_k, & \text{if } M_k|e_k| > 1. \end{cases} \tag{4}$$

The signal after demodulation is routed to the estimator unit of BS. The estimator unit calculates the current estimate $\hat{x}_k$ using the Kalman type equation (5) which determines whether the contribution of the previous estimate or the contribution of the current observation will be greater in calculating the current estimate based on the value of L_k, set by the estimator unit.

$$\hat{x}_k = \hat{x}_{k-1} + L_k \tilde{y}_k, \tag{5}$$

L_k controls how quickly estimates $\hat{x}_k$ converge to the transmitted sample's original value x_t. The value of L_k is calculated by the optimization task. The transmission unit receives the most recent estimate via the feedback channel. The transmission unit compares the received signal on the feedback channel to the original sample value and, the difference is sent to the modulator. The EM unit sets the value of L_k to L_{k+1} and modulation index M_k to M_{k+1} and the next $(k + 1)^{th}$ cycle begins. After n cycles, the final estimate $\hat{x}_n$ of the sample x_t is routed to the addressee and the next sample transmission begins. The initial values $\hat{B}_1$ and $\hat{M}_1$ are established by the saturation factor α which determines the allowable level of the likelihood of transmitter saturation, the variance of the input signal σ^2_0, and the mean value x0 of the input signal respectively.

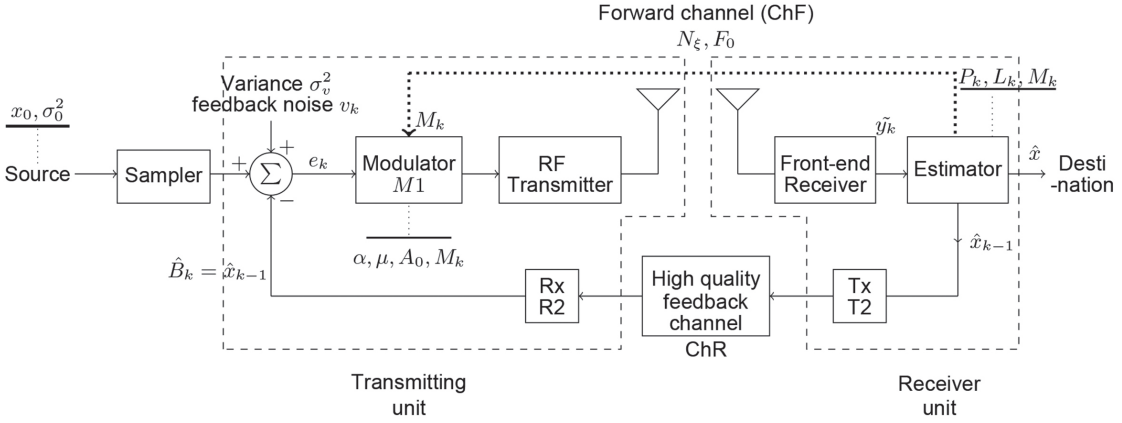


Fig. 1. System architecture of wireless Analog Feedback Communication

3. AFCS OPTIMIZATION FOR RAYLEIGH FADING CHANNEL

The MSE of current sample estimates can be calculated at the base station as the prior distribution of the input is assumed to be known at the receiver and the algorithm mentioned in section 2 provides the values of the current estimates. MSE in the k^{th} iteration denoted by P^k is given by:

$$P_k = E[(x - \hat{x}_k)^2]. \tag{6}$$

MSE depends not only on the estimation algorithm but also on the parameters B_k and M_k of the adaptive modulator at the transmitter.

Hence we need to set the values of L_k; B_k; M_k to achieve minimum MSE. This in turn results in the optimization of both the transmission and reception algorithm. However, due to mathematical complications created by the saturation form of the modulator characteristic (2), a direct solution to this optimization problem is unattainable. Although, if a new transmission quality parameter – the permitted probability of over-modulation i.e. the probability of the appearance of errors due to over-modulation is introduced, this problem can be solved.

Probability of over-modulation of the adaptive modulator denoted by Pr_k^{over} in each iteration is given by:

$$Pr_k^{over} = Pr(M_k e_k > 1 | \tilde{y}_1^{k-1}, B_{k-1}, M_{k-1}) < \mu. \quad (7)$$

Parameters (B_k; M_k) satisfying the inequality given in (7) form a permissible set of parameters Ω_k which guarantee the fulfillment of the condition $M_k|e_k| \leq 1$ with probability not smaller than 1-μ, where μ is the probability of modulator saturation and is the measure of errors occurring due to loss of information when the modulator saturates. Hence, the quantity 1-(1-μ)$^n \approx n\mu$ represents the frequency of appearance of errors caused by over-modulation at the time of sample transmission and is equivalent to bit-error rate (BER) in digital communication systems. Usually, values of between $10^{-12} \leq \mu < 10^{-4}$ is practically enough for the design. Considering this value of μ and (7), the parameters of adaptive modulator B_k; M_k are calculated. Such an adaptive modulator is known to be a statistically fitted adaptive modulator.

The expression for M_k as proposed in [21] is a result of consideration of this statistical fitting condition and is given by:

$$M_k = \frac{1}{\alpha\sqrt{P_{k-1} + \sigma_v^2}}, \quad (8)$$

here α is the saturation factor which takes into account the probability of over-modulation and the statistical fitting of the modulator. The relation between α and μ is given by:

$$\Phi(\alpha) = \frac{1}{\sqrt{2\pi}} \int_0^{\alpha} e^{\frac{-x^2}{2}} dx \geq \frac{1-\mu}{2}, \quad (9)$$

where $\Phi(\alpha)$ is known as Gaussian error function.

The statistically fitted modulator almost invariably operates as a linear unit, and the non-linear transmitter model (2) can be substituted by the linear one:

$$y_k = A_0 \hat{M}_k e_k = A_0 \hat{M}_k (x_k - B_k), \quad (10)$$

which in turn allows to replace the model in (4):

$$\tilde{y}_k = A_0 h \hat{M}_k (x_k - B_k) + \zeta_k. \quad (11)$$

The differences in the working of the model (2), (4) and the statistically fitted AFCS constructed using models, (10), (11) may emerge with a probability of $n\mu$. As a result, the MSE of the two systems might only vary by $O(n)$ order. However, the modulator's statistical fitting enables the transformation of a nonlinear optimization problem into a linear one. This task can then be solved using the Bayesian estimation theory approach [21],[22].

3.1. OPTIMIZATION OF RAYLEIGH FADING CHANNEL BASED ON MSE PERFORMANCE

As the fading coefficient is a complex quantity, the received signal and estimate signal given by (5) will also be complex and the MSE P_k takes the form:

$$P_k = E[(x_k - \hat{x}_k)(x_k - \hat{x}_k)^*]. \quad (12)$$

Substituting (5) into equation (12), we have

$$P_k = E[(x_k - \hat{x}_{k-1} - L_k \tilde{y}_k)(x_k - \hat{x}_{k-1} - L_k \tilde{y}_k)^*]. \quad (13)$$

By differentiating P_k w.r.t. L_k* and equating it to zero, we find the optimal value of L_k which minimizes the MSE:

$$L_k = \frac{E[(x_k - \hat{x}_{k-1})\tilde{y}_k^*]}{E[\tilde{y}_k \tilde{y}_k^*]}. \quad (14)$$

On substitution from (11), simplification and considering $E[(x_k - \hat{x}_{k-1})] = 0$, we get:

$$E[(x_k - \hat{x}_{k-1})\tilde{y}_k^*] = A_0 M_k h^* P_{k-1}, \quad (15)$$

where $P_{k-1} = E[(x_k - \hat{x}_{k-1})(x_k - \hat{x}_{k-1})^*]$. Now as $hh^* = |h|^2$, $E\{v_k v_k^*\} = \sigma_v^2$, $E\{\zeta_k \zeta_k^*\} = \sigma_\zeta^2$

$$E[\tilde{y}_k \tilde{y}_k^*] = |h|^2 A_0{}^2 M_k{}^2 (P_{k-1} + \sigma_v^2) + \sigma_\zeta^2. \quad (16)$$

Hence from (14),(15) and (16) L_k is given by

$$L_k = \frac{A_0 M_k h^* P_{k-1}}{\|h\|^2 A_0{}^2 M_k{}^2 (P_{k-1} + \sigma_v^2) + \sigma_\zeta^2}. \quad (17)$$

Substituting (17) into (13), with (11) and considering MSE in the previous cycle as given by:

$$P_{k-1} = E[(x_k - \hat{x}_{k-1})(x_k - \hat{x}_{k-1})^*]. \quad (18)$$

After simplification, we get:

$$P_k = \frac{[\|h\|^2 A_0{}^2 M_k{}^2 \sigma_v^2 + \sigma_\zeta^2] P_{k-1}}{\|h\|^2 A_0{}^2 M_k{}^2 (P_{k-1} + \sigma_v^2) + \sigma_\zeta^2}. \quad (19)$$

It is evident from the equation (19) that the MSE of transmission decreases with increasing modulation depth M_k and is independent of the control B_k values.

As MSE is independent of control values B_k, the optimization task is simplified to searching for the maximum value M_k from the set of values that satisfy the condition (7).

From equation (8), it seems that M_k values do not depend on the observations hence, they can be independent of the channel model used. Another control i.e. B_k is the estimate $\hat{x}_{k-1}$ during the previous cycle:

$$B_k = \hat{x}_{k-1}(\tilde{y}_1^{k-1}) = \hat{x}_{k-1}. \quad (20)$$

To summarize, the algorithm adopted by AFCS which ensures optimum transmission reception and achieves minimum mean square error is given by:

1. Calculate the estimate at the receiver using equations: (5) and (17)
2. Altering the transmitting unit's adaptive modulator in accordance with equations (20) and (8).

Initial values considered for the iterative algorithm: $\hat{x}_0 = x_0$; $P_0 = \sigma_0^2$.

3.2. DIVERSITY COMBINING WITH AFCS

Now consider that the output of the transmitter is transmitted using a single antenna but the receiver has multiple antennas. Fig. 2. shows the general case for N receive antennas. Let us consider N = 2.

The transmitted signal takes two independent channels $h1$ and $h2$. At the receiver we have:

$$\tilde{y}_{ki} = A_0 h_i \hat{M}_k (x_k - B_k) + \zeta_{ki}, i \in \{1, 2\}. \quad (21)$$

The received signals $\tilde{y}_{k1}$ and $\tilde{y}_{k2}$ can be combined at the receiver using one of the diversity combining techniques like: selection combining (SC), equal gain combining (EGC) or maximum ratio combining (MRC) [1]. In SC, simply the strongest link is selected for transmission. As a result, if $h1 > h2$, the combined sequence at the receiver is $\tilde{y}_{k(SC)} = \tilde{y}_{k1}$, otherwise $\tilde{y}_{k(SC)} = \tilde{y}_{k2}$ and the corresponding channel coefficient is $hSC = max(h_1; h_2)$, whereas the noise variance remains the same as σ^2_ζ.

EGC gives equal weightage to both the received sequences. Thus, $\tilde{y}_{k(EGC)} = \tilde{y}_{k1} + \tilde{y}_{k2}$ and $h_{EGC} = h_1 + h_2$ give the received sequence and equivalent channel coefficient respectively. Due to equally adding both received signals, the noise variance becomes $2\sigma^2_\zeta$.

Finally, MRC blends the two sequences weighted by a fading coefficient factor. MRC is known to outperform SC and EGC and is known to be an optimum combining technique. The SNR improvement offered by MRC is better compared to SD and EGC [23].

3.3. MRC DIVERSITY COMBINING WITH AFCS

We find the effect of having receiver diversity (two receiving antennas) and MRC combining at the receiver on the equations and algorithm mentioned in section (3).

The received signal after MRC combining is given by:

$$\tilde{y}_{k(MRC)} = {w_1}^* \tilde{y}_{k1} + {w_2}^* \tilde{y}_{k2}, \quad (22)$$

where $w_1 = \frac{h_1}{\sqrt{|h_1|^2+|h_2|^2}}$ and $w_2 = \frac{h_2}{\sqrt{|h_1|^2+|h_2|^2}}$ are weights assigned to each channel (h_1 and h_2). The signal received on channel with fading coefficient h_1 is $\tilde{y}_{k1} = h_1 A_0 M_k (x_k - \hat{x}_{k-1} + v_k) + \zeta_{1k}$ and signal received on channel with fading coefficient h_2 is $\tilde{y}_{k2} = h_2 A_0 M_k (x_k - \hat{x}_{k-1} + v_k) + \zeta_{2k}$. Here ζ_{1k} and ζ_{2k} are uncorrelated AWG noises in the two diversity channels. Now this received and combined signal will be given to the estimator which calculates the estimate same as in section (2) equation (5).

$$\hat{x}_k = \hat{x}_{k-1} + L_k \tilde{y}_{k(MRC)}. \quad (23)$$

Similar to how the expressions of L_k and P_k were evaluated in section (3), new expressions for L_k and P_k will be evaluated with the received and combined signal $\tilde{y}_k(MRC)$. Hence substituting $\tilde{y}_k(MRC)$ in equation (14) in place of $\tilde{y}_k$, we get:

$$L_k = \frac{E[(x_k - \hat{x}_{k-1})\tilde{y}^*_{k(MRC)}]}{E[\tilde{y}_{k(MRC)}\tilde{y}^*_{k(MRC)}]}. \quad (24)$$

From equation (22) and following the same steps as in section 3.1 we get:

$$L_k = \frac{A_0 M_k P_{k-1}(w_1 h_1^* + w_2 h_2^*)}{E[\tilde{y}_{k(MRC)}\tilde{y}^*_{k(MRC)}]}, \quad (25)$$

where denominator term is given by:

$$\begin{aligned} &E[\tilde{y}_{k(MRC)}\tilde{y}^*_{k(MRC)}] \\ &= \sigma^2_\zeta(|w_1|^2 + |w_2|^2) \\ &+ {A_0}^2 {M_k}^2 (P_{k-1} + \sigma^2_v)(|w_1|^2|h_1|^2 + |w_2|^2|h_2|^2 \\ &+ w_1^* h_1 w_2 h_2^* + w_1^* h_1 w_2^* h_2), \end{aligned} \quad (26)$$

and,

$$\begin{aligned} &P_k = \\ &P_{k-1} - {L_k}^* A_0 M_k P_{k-1}(w_1 h_1^* + w_2 h_2^*) \\ &- L_k A_0 M_k P_{k-1}(w_1^* h_1 + w_2^* h_2) \\ &+ |L_k|^2[\sigma^2_\zeta(|w_1|^2 + |w_2|^2) \\ &+ {A_0}^2 {M_k}^2 (P_{k-1} + \sigma^2_v)(|w_1|^2|h_1|^2 + |w_2|^2|h_2|^2 \\ &+ {w_1}^* h_1 w_2 h_2^* + w_1 h_1^* w_2^* h_2)]. \end{aligned} \quad (27)$$

4. PERFORMANCE EVALUATION

The performance of a communication channel in general depends on two factors: the properties of the channel and the properties of the source. When the smallest number of bits required to represent the source equals the highest number of bits possible on the channel, we reach the theoretically achievable the best performance limit (OPTA). OTPA can be calculated by finding the distortion rate function $DRF = D(R = C)$ at a rate equal to channel capacity. It may appear simple to evaluate the DRF at a rate equal to the channel capacity; however, the issue is that the DRF rate is defined in bits per source sample, whereas the channel capacity is specified in bits per channel use. Since in AFCS, we are transmitting one source sample in n number of cycles i.e. n, channel uses, the capacity must be scaled accordingly. According to rate distortion theory, [24, 25], the output signal to distortion ratio (SDR) for a Gaussian source is given by:

$$\begin{aligned} D &= 2^{-2R}\sigma_x^2, \\ SDR &= \left(\frac{\sigma_x^2}{D}\right) = 2^{2R}. \end{aligned} \quad (28)$$

The OTPA equals:

$$OTPA = D(R)|_{R=KC/N} = D\left(\frac{KC}{N}\right). \quad (29)$$

In the case of AFCS, we are sending one sample in n cycles hence $K = n$ and $N = 1$, so the OTPA becomes:

$$OTPA = D(R)|_{R=nC} = D(nC). \quad (30)$$

The capacity of AFCS with a slow fading channel is found next.

4.2. CHANNEL CAPACITY

For a flat fading channel model with perfect channel knowledge at the receiver, the capacity of forward single input single output (SISO) channel is given by [1, 26]:

$$\begin{aligned} C &= \log_2(1 + SNR) \\ &= \log_2\left(1 + \frac{A_0^2 M_k^2 |h|^2 (P_{k-1} + \sigma_v^2)}{\sigma_\zeta^2}\right), \end{aligned} \quad (31)$$

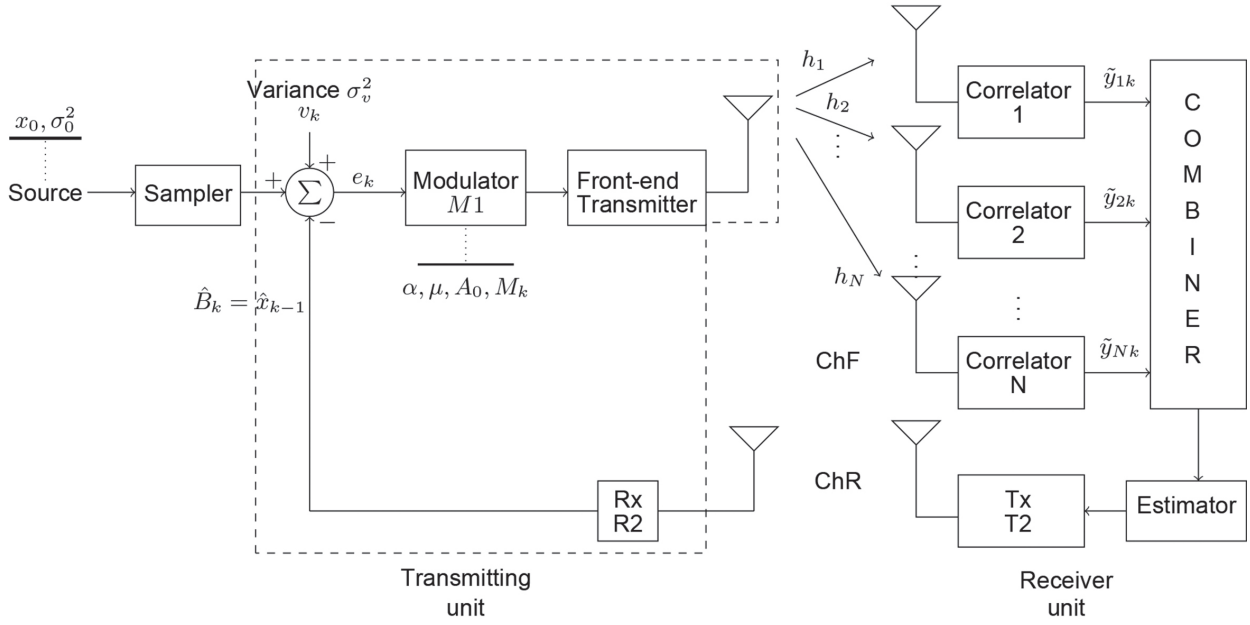


Fig. 2. System architecture of wireless AFCS with receive diversity

Here capacity is expressed in *bits/s/Hz*. Here h is considered to be a flat fading complex channel impulse response of block fading type. Even though h is constant for a block of transmitted symbols but as it is random, the channel capacity also becomes random. Hence ergodic capacity and outage probability make more sense in the case of fading channels. The ergodic capacity is defined as the statistical average of the mutual information where the expectation is taken over $|h|^2$ [26, 23] and is given by:

$$C_{erg} = E\left\{ \log_2 \left(1 + \frac{A_0^2 M_k^2 |h|^2 (P_{k-1} + \sigma_v^2)}{\sigma_\zeta^2} \right) \right\}. \quad (32)$$

By Jensen's inequality [1] applied to (32):

$$\begin{aligned} &E\left\{ \log_2 \left(1 + \frac{A_0^2 M_k^2 |h|^2 (P_{k-1} + \sigma_v^2)}{\sigma_\zeta^2} \right) \right\} \\ &\leq \log_2 \left(1 + \frac{A_0^2 M_k^2 E\{|h|^2\} (P_{k-1} + \sigma_v^2)}{\sigma_\zeta^2} \right), \end{aligned} \quad (33)$$

where E{.} stands for expectation. Finally, the OTPA considering optimum AWGN capacity and DRF given by (28) is:

$$\begin{aligned} SDR &= \left(\frac{\sigma_x^2}{D} \right) = 2^{2nC} \\ SDR_{dB} &= \left(\frac{\sigma_x^2}{D} \right)_{dB} = 20n \log_{10} \left(1 + Q_F^2 \right) \end{aligned} \quad (34)$$

5. RESULTS AND DISCUSSION

Based on the results found in section 3 and section 4, the algorithm pertaining to the Rayleigh fading channel model is implemented and simulation is performed in MATLAB. AFCS in Rayleigh fading channel with single input single output (SISO) without diversity is compared with the system with receive diversity with N number of receive antenna ($N = 2; 4; 6$) is investigated and the results are presented.

For simulation, a random input signal with mean $x_0 = 1$, variance $\sigma^2_0 = 0.625$, and band-limited to $2.5kHz$ is taken into account which is sampled at a frequency of $8kHz$. The plots are shown in Fig. 3 to Fig. 6.

MSE between the input and estimate is considered to be the overall performance criteria. Fig. 3 shows MSE per iteration for SISO Rayleigh fading channel, and SIMO Rayleigh fading channel with MRC for 2,4, and 6 receiving antennas. Convergence and MSE performance comparison for two different channel signal-to-noise Ratio (SNR) values $5dB$ and $10dB$ can be observed in Fig. 3a and 3b respectively. MSE performance is better for 10dB for obvious reason of higher channel SNR. The evident improvement with the inclusion of receive diversity can be seen in both cases. For all the schemes the MSE for the proposed communication system architecture (Analog Feedback Communication System) converges to zero. The values of MSE in SISO case during the first iteration is higher compared to SIMO case, and it decreases as the number of diversity paths increases. Furthermore, With an increase in the number of receive antennas, the MSE per iteration decreases, and the convergence to zero MSE happens in fewer iterations compared to SISO Rayleigh fading channel scheme. Fig. 4 shows the plot showing the variation in AFCS output SNR for different channel SNR. It shows an evident improvement in SNR at channel output with receive diversity. The improvement in SNR is about 15dB at 20 dB channel SNR in the case of 6 receive antennas compared to the SISO channel. This improvement in SNR certainly leads to the improvement in approaching the capacity limit. Fig. 5 shows the maximum achievable capacity limits

for AWGN AFCS compared to Rayleigh fading AFCS. The graph is plotted using (32). The maximum achievable performance limit of Rayleigh fading AFCS is less than that of AWGN. The gap can be made small by using spatial diversity and MIMO techniques. Spatial diversity in terms of multiple receive antennae to mitigate the effect of fading channels is the subject of the current paper.

Finally, in Fig. 6a, the behavior of AFCS with Rayleigh fading channel and receive diversity is shown on the classic spectral efficiency versus bit energy to noise ratio plot for ideal systems. Here the operating points of M-ary PSK digital communication system are shown and are compared with the operating points of diversity-enabled AFCS with N=1,2,4,6 receive antennas. Diversity-enabled AFCS seems to coincide with ideal systems which achieve Shannon's capacity. The zoomed version of the plot in Fig. 6b shows that the spectral efficiency improves with an increase in the number of diversity branches.

5.2. COMPARISON AND DISCUSSION

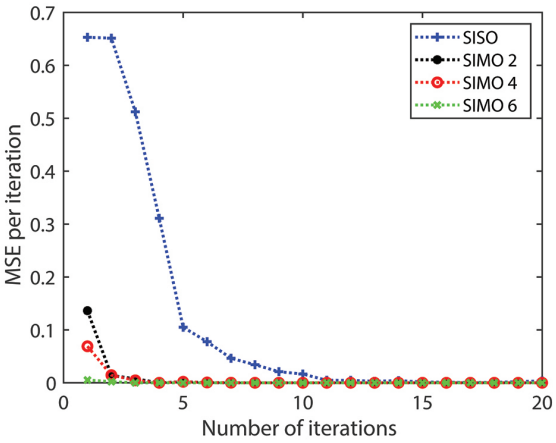
We proposed an analog communication system with the Rayleigh fading channel model in this paper. A one-to-one comparison with the current wide range of digital communication systems with advanced coding and decoding techniques does not make sense because the majority of the signal processing chain, and performance criteria focused is different in AFCS and DCS.

In AFCS MSE is the performance criteria and both transmitter and receiver are optimized to obtain minimum MSE as shown in section 3 Whereas in digital communication systems (DCS), there are multiple performance criteria that are correlated like bit-rate, bit-error-rate (BER), power-bandwidth efficiency, etc. The search for an optimum balance between these performance criteria involves tradeoffs. For example, increasing the bit rate decreases the BER, to achieve a good BER more power must be spent, and so on. Many stages are involved in DCS like digitization, source-channel coding/decoding, modulation, etc., and optimizing each of these processes for common performance criteria is practically very difficult and leads to a very complex, power-consuming system.

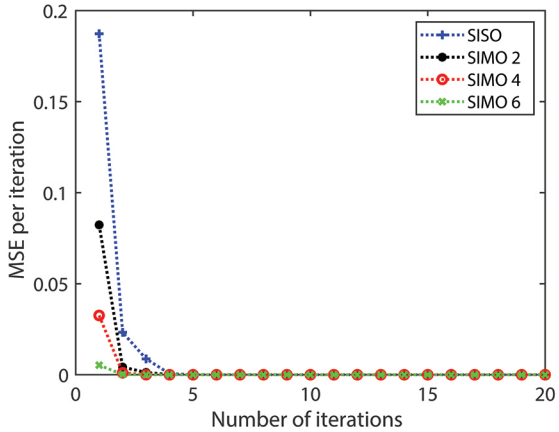
The AFCS uses feedback and iterative estimation algorithm instead of complex coding. Hence will be a better choice for applications that are power constrained such as wireless sensor networks, satellite communication, and many applications where battery-operated sensors transmit information to base stations. In applications where source information is analog, converting it to digital and losing the information in the process of quantization, incorporating complex coding and decoding algorithms to reach the performance limits seems a costly and complicated process. Instead, if we transmit uncoded analog information directly, and utilize the feedback which is available in almost all scenarios of communication these days, we can utilize the available communication resources efficiently. There is definitely an iterative process and feedback involved, but as shown in Fig. 3, with the diversity scheme, it takes hardly 2-3 iterations to reach close to zero mean square error (MSE). The mean square error quickly converges to zero as we increase the number of antennae (N).

In DCS a very low BER does not ensure that an analog signal will be transmitted with reliability. The quantization loss always will be present. Whereas in the case of AFCS transmission a very low or almost zero MMSE ensures reliable transmission.

The design of digital communication stages is done keeping in mind a particular channel SNR as a design parameter. If the channel conditions are poorer compared to that, then the performance of DCS drastically deteriorates but does not improve if the channel condition is better than the design parameter. This is known as the cliff effect, and DCS suffers from the cliff effect. In AFCS symbols are transmitted hence there is no problem of delay which occurs when long block lengths are used. The adaptive modulator is adjusted according to the channel conditions iteratively, hence there is no cliff effect problem and there is a graceful degradation in case of poor channel conditions.



(a) MSE per iteration for SNR 5dB



(b) MSE per iteration for SNR 10dB

Fig. 3. MSE per iteration for SISO Rayleigh fading channel, SIMO Rayleigh fading channel with MRC for 2 ,4 and 6 receiving antennas.

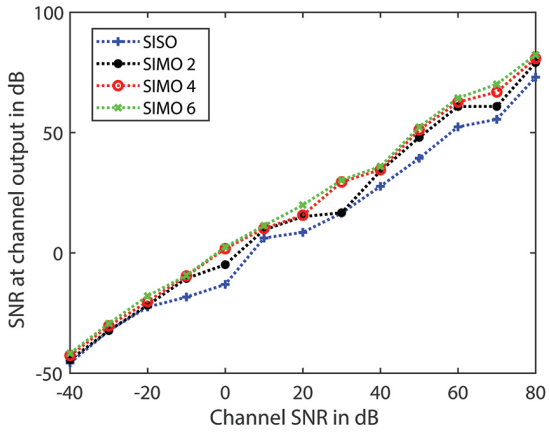


Fig. 4. Graph showing the variation in AFCS output SNR for different channel SNR

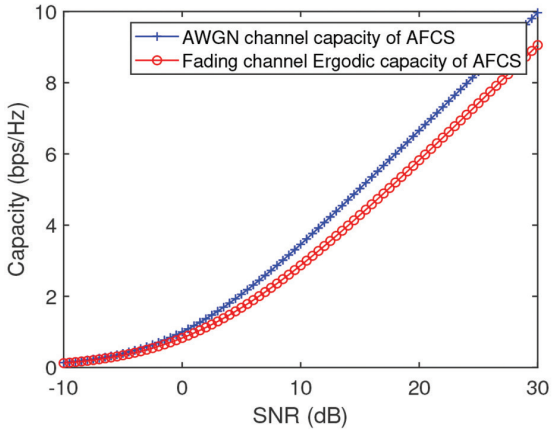
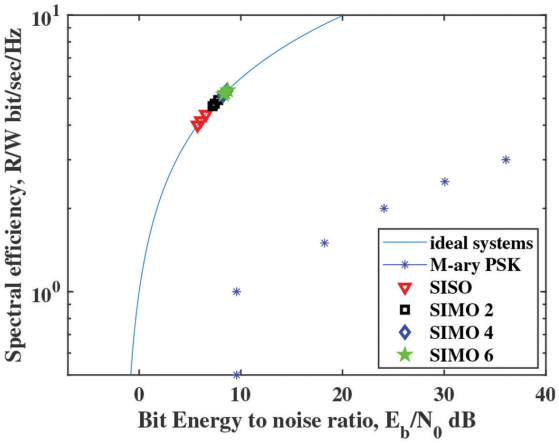
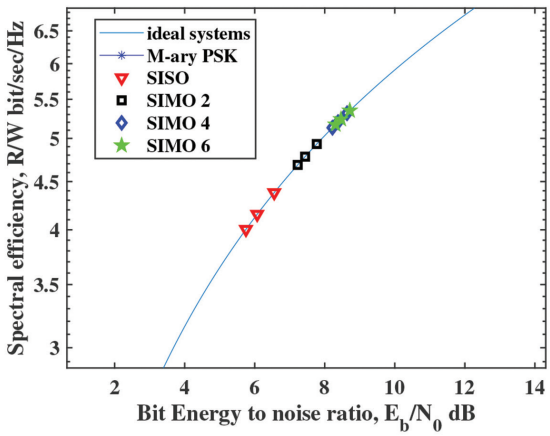


Fig. 5. Graph showing the comparison of maximum achievable capacity, in case of an AWGN AFCS channel and AFCS with fading channel model



(a) Spectral efficiency versus bit energy to noise ratio plot (normal scale)



(b) Zoomed in version

Fig. 6. Spectral efficiency versus bit energy to noise ratio

6. CONCLUSION

A system model is created and a thorough analysis is provided for AFCS with Rayleigh fading channel model and MRC receive diversity. It can be seen that the system can reach the Shannon capacity boundary which is something that most spectrally efficient digital communication systems cannot achieve. Because the AFCS system does not involve digitizing and coding, many of the procedures such as analog-to-digital conversion, encoding, digital-to-analog conversion, decoding, and so on are not required for AFCS. This leads to significant savings of energy which in turn results in the transmitter becoming lighter and more power efficient. Receive diversity used with MRC in AFCS results in enhanced MSE performance and improved power bandwidth efficiency. OTPA performance is also seen to improve with diversity. Since power requirements for short- to medium-range applications involving sensor networks are relatively strict, such a system could be effectively used in those applications. By using diversity strategies for AFCS, SNR and coverage area can both be further increased.

We are investigating more diversity strategies including MIMO channels in AFCS as future work. MIMO AFCS is anticipated to produce better outcomes by mitigating the effects of fading channels.

7. REFERENCES

[1] A. Goldsmith, Wireless communications. Cambridge university press, 2005.

[2] C. Shannon, "A mathematical theory of communication, part i," The Bell System Technical Journal, vol. 27, pp. 379–423, 1948.

[3] C. Shannon, "The zero error capacity of a noisy channel," IRE Transactions on Information Theory, vol. 2, no. 3, pp. 8–19, 1956.

[4] C. Shannon, "Communication in the presence of noise," Proceedings of the IRE, vol. 37, no. 1, pp. 10–21, 1949.

[5] C. Berrou and A. Glavieux, "Near optimum error correcting coding and decoding: Turbo-codes," IEEE Transactions on communications, vol. 44, no. 10, pp. 1261–1271, 1996.

[6] J. Hagenauer, E. Offer, and L. Papke, "Iterative decoding of binary block and convolutional codes," IEEE Transactions on information theory, vol. 42, no. 2, pp. 429–445, 1996.

[7] B. Sklar et al., Digital communications, vol. 2. Prentice hall Upper Saddle River, NJ, USA:, 2001.

[8] T. Goblick, "Theoretical limitations on the transmission of data from analog sources," IEEE Transactions on Information Theory, vol. 11, no. 4, pp. 558–567, 1965.

[9] P. Elias, "Channel capacity without coding, quarterly progress report, research laboratory of electronics," 1956.

[10] T. Kailath, "An application of shannon's ratedistortion theory to analog communication over feedback channels," Proceedings of the IEEE, vol. 55, no. 6, pp. 1102–1103, 1967.

[11] J. Omura, "Optimum linear transmission of analog data for channels with feedback," IEEE Transactions on Information Theory, vol. 14, no. 1, pp. 38–43, 1968.

[12] A. Platonov and I. Zaitsev, "New approach to improvement and measurement of the performance of phy layer links of wsn," IEEE Transactions on Instrumentation and Measurement, vol. 63, no. 11, pp. 2539–2547, 2014.

[13] A. Platonov and I. Zaitsev, "Perfect low power narrowband transmitters for dense wireless sensor networks," in Proceedings of the Signal Processing: Algorithms, Architectures, Arrangements, and Applications, pp. 332–337, 2018.

[14] A. Platonov, "Energy saving adaptive sdr transmitters for small satellites and their metrological characteristics," in Proceedings of the 5th IEEE International Workshop on Metrology for AeroSpace (MetroAeroSpace), pp. 468–472, 2018.

[15] A. Platonov, "Theoretical preconditions for development of perfect communication channels" small satellites–earth"," in Proceedings of the IEEE 5th International Workshop on Metrology for AeroSpace (MetroAeroSpace), pp. 142–147, 2019.

[16] A. Platonov, "Perfect signal transmission using adaptive modulation and feedback," in Modulation in Electronics and Telecommunications, p. 1, IntechOpen, 2020.

[17] A. Platonov, I. Zaitsev, and L. Opalski, "Theoretical basis, principles of design, and experimental study of the prototype of perfect afcs transmitting signals without coding," in Proceedings of Photonics Applications in Astronomy, Communications, Industry, and High Energy Physics Experiments 2017, vol. 10445, pp. 520–533, 2017.

[18] A. Platonov, "Bayesian methods in optimization of analog communication systems with feedback," Materiały X Poznańskich Warsztatów Telekom., PWT-2012, Poznań, pp. 32–35, 2012.

[19] A. Platonov, "Transmission of signals from the analog sources: criterions in design of top performance acs," Przegląd Telekomunikacyjny, vol. 4, pp. 380–383, 2015.

[20] A. Platonov, "Capacity and power-bandwidth efficiency of wireless adaptive feedback communication systems," IEEE Communications Letters, vol. 16, no. 5, pp. 573–576, 2012.

[21] A. A. Platonov, "Optimization of adaptive communication systems with feedback channels," in Proceedings of the IEEE Wireless Communications and Networking Conference, pp. 1–6, 2009.

[22] H. L. Van Trees, Detection, estimation, and modulation theory, part I: detection, estimation, and linear modulation theory. John Wiley & Sons, 2004.

[23] M. Viswanathan, Wireless Communication Systems in Matlab. Independent, 2020. [24] N. Goertz, Joint source-channel coding of discrete-time signals with continuous amplitudes, vol. 1. 2007.

[25] T. M. Cover, Elements of information theory. John Wiley & Sons, 1999.

[26] D. Tse and P. Viswanath, Fundamentals of wireless communication. Cambridge university press, 2005.

Designing a W-type index praseodymium-doped chalcogenide fiber for strongly-efficient MIR Laser beyond 4 μm

Original Scientific Paper

M A Khamis
Middle Technical University,
Baquba Technical Institute, Electrical Techniques Department
Muasker Al Rashid Street, Baghdad, Iraq
mustafaelectronic@mtu.edu.iq

Ziad QaisAlAbbasi
Middle Technical University,
Baquba Technical Institute, Electrical Techniques Department
Muasker Al Rashid Street, Baghdad, Iraq
ziad.al-abbasi@mtu.edu.iq

Abstract *– This study is to propose an efficient design based on W-type index praseodymium (Pr^{3+})-doped chalcogenide fiber to enhance the emission of mid-infrared (MIR) wavelength beyond 4 μm. The advantage of the proposed design is to enlarge the core diameter to 32 μm under single mode guidance in order to restrict any impacts of nonlinearity and enhance the applied pumping power before the damage threshold. The effective mode field areas are about 979 μm^2 and 812 μm^2 at 4.5μm emitted MIR wavelength and 2.04μm pump wavelength, respectively. In the considered laser layout within this paper, a single pair of Fiber Bragg Gratings FBGs lay is adopted in the overlapping area of the emitting cross-sections parts of Pr^{3+} ions in the transitions ($^3F_2$, $^3H_6 \rightarrow {}^3H_5$ and $^3H_5 \rightarrow {}^3H_4$). This selected laser layout avoids the difficulties in the fabrication of concatenated FBGs in Pr^{3+}-doped chalcogenide glass fibers. In addition, the efficiency of the laser can be also enhanced at this laser layout by emitting two photons in MIR wavelengths from one excited ion. The simulation outcomes indicate the possibility to exceed 61% of slope efficiency at 4.5 μm for 1dB/m loss in the fiber.*

Keywords*: mid-infrared laser, chalcogenide glass material, praseodymium-doped fiber,W-index structure*

1. INTRODUCTION

Recently, mid infrared (MIR) light sources beyond 4μm have drawn considerable interests for applications in many industrial and scientific applications, including medical diagnosis, gas sensing, atmosphere, defence, food security, environment monitoring, etc [1-4]. This is because their wavelength region has strong absorption strength of radiations, and covers the main atmospheric windows as well as many molecular gases, soil pollutants, toxic agents, water and breath components [5]. MIR wavelength range can be generated from several light sources such as optical parametric amplifiers [6], quantum cascade laser [7], supercontinuum [8] and a fiber laser [9-11].

Among others light sources, lanthanide ion doped fiber lasers are promising light sources in the MIR fluorescence between 3-5μm wavelength range. Praseodymium (Pr^{3+}) has attractive characteristics in the MIR fluorescence among others lanthanide ions [12]. This is due to an available pumping of Pr^{+3} from 2 μm thulium diode fiber laser. In addition, the pumped absorption cross-section of Pr^{3+} is high as compared to other lanthanide ions. Furthermore, the overlapping emission source is represented by a cross sectional area that lies between two levels. The first level is $^3F_2$, $^3H_6 \rightarrow {}^3H_5$ (3.3–4.7 μm), on the other hand, the second level is $^3H_5 \rightarrow {}^3H_4$ (3.7–5.5 μm). Which yields to avoid the need of the cascade lasing approach and enhance the MIR emission. Finally, Pr^{3+} has a strong cross relaxation process, leading to excite three ions to manifold $^3H_5$ by de-excitation of one ion to $^3F_3$ manifold [13]. However, it is prerequisite to use host materials with low phonon energy and optical loss to release an efficient MIR light source from a Pr^{3+}- doped fiber and the promising candidate for this purpose is the chalcogenide glass [14]. This host material has attractive properties such as good rare earth ion solubility and high refractive index. Also, the light propagates through this glass material for a broad range of wavelengths (2-14μm). However, the main

problem to release MIR laser from chalcogenide fiber is the low damage threshold in which limits the applied pump power [15].

For this purpose it is required to make the fiber core with a large mode field area to increase the possible value of the applied pumping power before reaching its damage threshold. Unfortunately, this way has a limitation due to the enlarging of the core diameter which in turn drives the conventional fiber to multimode guidance that degrades the beam quality. Another way to overcome this limitation is by applying double cladding pumping scheme [9]. Despite enlarging the mode field area of the pump, the total laser efficiency is still limited to only 18% due to a low confinement factor of the pump at this scheme. Another possible way to enlarge the mode area of Pr^{3+}-doped chalcogenide is based on a photonic crystal fiber PCF [16]. Although, this fiber design offers more flexibility and optical characteristics than the conventional fiber but the fabrication and fused splices represent more challenges.

A W-type fiber is an alternative approach to overcome the above limitation and to enlarge the core area. This fiber profile has potential characteristics such as a dispersion-flattened fiber. In addition, it has a tight confinement mode, and as it evaluated aroundV = 3.8, it is possible to say it has alarge V-parameter as compared to the value V = 2.405 in the conventional fiber [8]. Having a large-valued V-parameter means it is practically possible to widen the fiber core diameters and, at the same time, maintaining the fiber operation to be in the single operation mode. In addition, it is easy to splice this fiber type with the standard step index fiber (SIF) and hence it becomes appropriate for all fiber laser systems. The W-type index structure has been already considered for ytterbium-doped fiber laser operating at 1.077μm [17], erbium-doped fiber amplifier at S-band [18], Neodymium-doped fiber laser at 914nm [19], and a bismuth-doped germanosilicate fiber laser in E band [20].

This study adopts Pr^{3+}-doped chalcogenide fiber laser based on W-type index structure is designed for MIR emission beyond 4μm. The core of the proposed design has a diameter of around 32 μm. In addition, the field area of the considered effective mode is about 979 μm^2 at 4.5μm laser wavelength, and it is valued as 812 μm^2 at 2.04 μm pump wavelength while maintained intrinsically single mode operation. For the laser scheme, a cascade pumping scheme is avoided by operation in the overlapping transition wavelengths of a Pr^{3+}-doped chalcogenide fiber where a single pair Fiber Brag Grating is selected at 4.5μm wavelength. Our results reflect that the laser power efficiency can be theoretically reached to 61% at 4.5μm of MIR emission with 1dB/m of fiber loss.

2. DESIGN AND THEORY

The proposed fiber design consists of three material layers with different refractive indices. The largest and lowest refractive indices are the core and the inner cladding, respectively as shown in Fig. (1). The core layer is composed of Pr^{3+}: GeAsGaSe glass material (n_1), whilst the outer cladding region consists of GeAsGaSe-S glass material (n_3). A suitable percentage around (0.8 atomic %) of Sulphur (S) is added instead of Selenium (Se) in the outer cladding layer to reduce the refractive index [21]. The inner cladding layer is also composed of GeAsGaSe-S glass material (n_2) but with twice percentage of addition Sulphur instead of Selenium to that in the outer cladding to make the refractive index of this region the lowest one. Fig. 2 indicates the refractive indices for several aspects of the fiber against the wavelength value. In particular it includes the core, the inner cladding and outer cladding against various wavelength levels.

By considering the W-index fiber, the normalized frequency (V) can be expresses as [8]:

$$V = \frac{2\pi a}{\lambda} \times NA = \frac{2\pi a}{\lambda}\sqrt{n_1^2 - n_3^2} \quad (1)$$

Where a is the radius of the core region, λ is the operating wavelength, NA is the numerical aperture, n1 and n3 are the refractive indices of the core and outer cladding, respectively. This parameter is important tolocate the cutoff wavelength of the proposed system where the operating wavelength must be higher than the specified cutoff wavelength. The W index fiber has large value of V-parameter about (3.8) under the condition of single mode. The advantage of this large value is to enlarge the core diameter and therefore the mode field area of the proposed fiber under single mode operation. In our proposed fiber, the core diameter is optimized to achieve the condition that the calculated V-parameter is lower than the value of 3.8 at single mode operation including the operating wavelengths range for the pump 2.04μm and laser emission 4.5μm. Fig. 3 shows the calculated V-parameter for both pump and laser wavelengths. The results indicate that 16 μm is the larger core radius to keep the operation under single mode in the chosen wavelengths for the proposed W-index fiber.

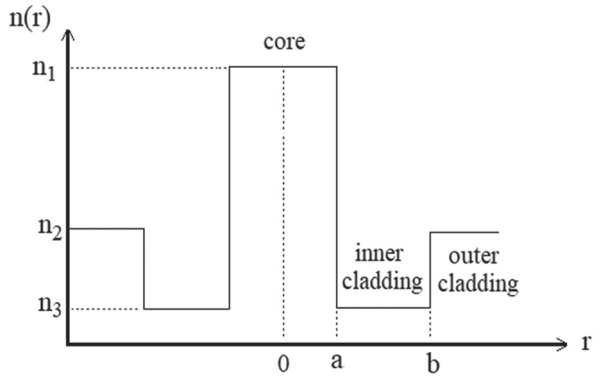


Fig. 1. Parameters structure of W-type fiber

A finite element method FEM that basison the idea of edge element is adopted in this paperfor the computation of the guidance properties for the W-type

Pr3+-doped chalcogenide fiber laser [16].The main advantages of applying FEM are to offer a mathematically elegant and rigorous formalism for modeling of complex geometrical shapes. In addition, a high degree of accuracy can be achieved with relatively short computation times. The propagation constant at the wavelengths 2.04 μm and 4.5 μm is calculated with three different refined meshes; with each newly generated mesh refinement possess a size that is half of the later one. Table (1) shows the computed convergence for the propagation factor with the mesh refinement approach. In the next simulation, the value of the propagation constant is obtained from the second mesh so as to provides the necessary level of accuracy for the subsequent calculations.

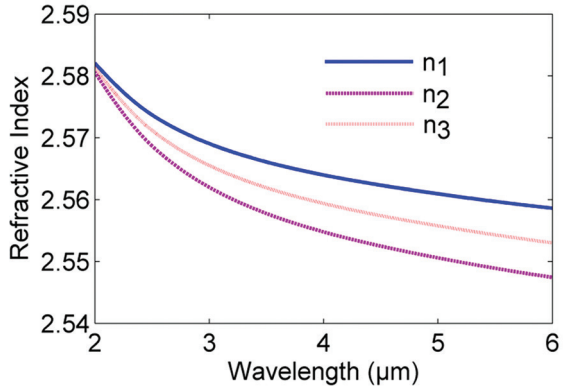


Fig. 2. The indices of refractionthat belongs to the core, and both the inner cladding and the outer cladding against various wavelength levels.

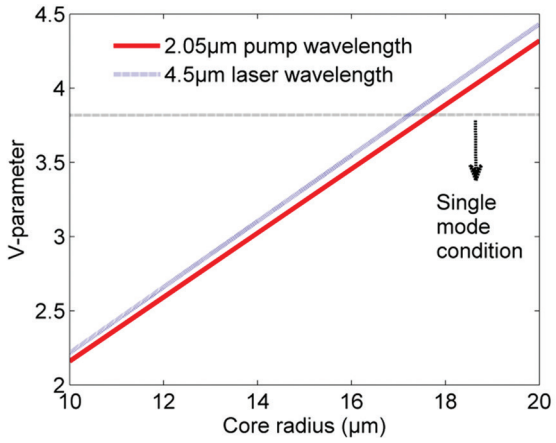


Fig. 3. The V-parameter calculated for the fiber of W-type at the pump and the laser wavelength with different core radius.

Table 1. The calculated propagation constant at pump wavelength 2.04μm and laser wavelength 4.5μm with two refined element meshes.

Mesh number	Propagation constant at 2.04μm.	Propagation constant at 4.5μm
280	7.948994 ×10⁶	3.575330×10⁶
1120	7.949039×10⁶	3.575435×10⁶
4480	7.949037×10⁶	3.575434×10⁶

At the fundamental mode, the electric field distribution is calculated at the pump wavelength 2.04μm and laser emission 4.5μm. Fig. 4 shows the distribution of the electric norm at 4.5μm wavelength, the confinement of the optical field as clearly indicated in the doped area of the fiber.

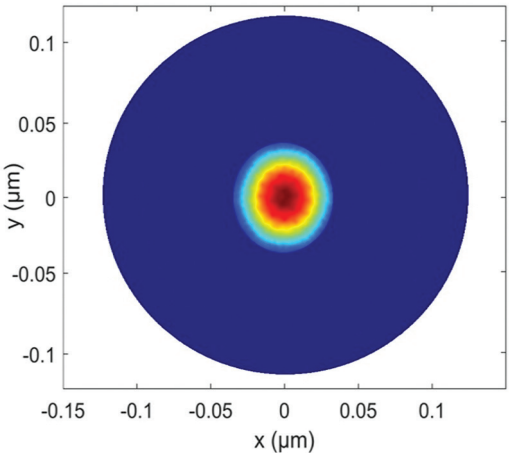


Fig. 4. The norm distribution depictedin fundamental mode of the electric field, wherethewavelength isλs= 4.5μm.

These calculated values of the electric field distribution are then exploited in the calculations of the field area of the effective mode as well as to obtain the factors of confinement at the operating wavelengths by applying the following expression [16]:

$$\Gamma_v = \frac{\iint_0^r \left|E_v^2(x,y)\right| dxdy}{\iint_0^\infty \left|E_v^2(x,y)\right| dxdy} \quad (2)$$

$$A_{eff} = \frac{\left(\iint_{-\infty}^{\infty} \left|E_v(x,y)\right|^2 dxdy\right)^2}{\iint_{-\infty}^{\infty} \left|E_v(x,y)\right|^4 dxdy} \quad (3)$$

Here v is to be substituted by either s for laser signal or p for pump, the radius doped area is denoted as r, the distribution of the electric field is represented by $E_v(x,y)$ at a given signal and a specified pumping wavelengths.In our design of the W-type indexed fiber, the field area of the effective mode and the confinement factor are about 979 μm2 and 0.63, respectively for the 4.5 μm wavelength. Moreover, they are valued about;812 μm^2 and 0.74, respectively for a wavelength of about 2.04 μm.

3. MODELLING OF THE FIBER LASER

In the case of a pump beam at 2.04 μm wavelength Pr^{3+}-doped chalcogenide ions can be represented as three energy levels system [9]. Fig. 5 indicates the main ions interaction of Pr^{3+}-doped chalcogenide at a low Pr^{3+} concentration level. There are two different emission transition bands relevant to Pr^{3+}-ions; namely, signal (1) (start-

ing from 3.3 up to 4.7 μm) between the manifolds ($^3F_2$, $^3H_6$) → $^3H_5$ and signal (2) (3.7–5.6 μm) between the manifolds $^3H_5$→$^3H_4$. As mentioned earlier in this article, there is an overlapping emission cross section between these two transition bands. In this study, the operating wavelength of the laser emission is chosen to be at 4.5 μm which is located in the overlapping region of the emission transition (starting from 3.7 up to 4.7 μm). This way assists in avoiding the need of the cascade lasing approach in which is merely one pair of FBGsthat would be required at 4.5 μm.

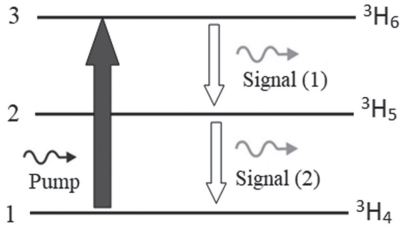


Fig. 5. Simplified energy level diagram for Pr^{3+}

The population inversion of the energy levels can be computed from solving the rate equations model of Pr^{3+}-doped chalcogenide fiber laser. During the steady state, these equations can be described in the following algebraic equations form [9,12]:

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} N_1 \\ N_2 \\ N_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ N \end{bmatrix} \quad (4)$$

The equation coefficients can be expressed as follow:

$$a_{11} = w_{13}^{p} \; ; \; a_{12} = w_{23}^{s_1} \; ; a_{13} = -w_{31}^{p} - w_{32}^{s_1} - \frac{1}{\tau_3}$$

$$a_{21} = w_{12}^{s_2} \; ; \; a_{22} = -w_{21}^{s_2} - w_{23}^{s_1} - \frac{1}{\tau_2} \; ; a_{23} = w_{32}^{s_1} + \frac{\beta_{32}}{\tau_3}$$

Here τ_2 and τ_3 are the lifetime of the manifolds $^3H_5$ and $^3H_6$, respectively; the transition branching ratio between $^3H_6$ to $^3H_5$ is denoted as β_{32}; w_{13}^{p} and $w_{23}^{s_1}$ are the pump stimulated rates; $_{23}$ and $w_{32}^{s_1}$ are the signal (1) stimulated rates; $w_{12}^{s_2}$ and $w_{21}^{s_1}$ are the signal (2) stimulated rates. These rates can be expressed as:

$$w_{ij}^{x} = \frac{\Gamma_x \sigma_{ij} \lambda_x}{Ahc} \times p_x \quad (5)$$

Here Γ_x and A refer to the confinement factor and effective mode area, respectively; x is either pump or signal (1) or signal (2); i and j are level transitions; σ_{ij} is the emission and absorption cross sections for specific ij transition. Finally, P_x is either substituted by the evolution power for pump or signal (1), otherwise it represents signal (2). Due to the operating wavelength at the overlapping region, so the power of signal (1) is the same at the signal (2). The spatial evolution of the pump P_p and signal P_s for all of the considered fiber length could be obtained from the solution of the following differential equations [12]:

$$\frac{dP_p^{\pm}}{dz} = \mp\Gamma_p \left[\sigma_{pa} N_1 - \sigma_{pe} N_3 \right] P_p^{\pm} \mp \alpha P_p^{\pm} \quad (3)$$

$$\frac{dP_s^{\pm}}{dz} = \mp\Gamma_s \left[\left(\sigma_{ia} N_2 - \sigma_{ie} N_3 \right) + \left(\sigma_{sa} N_1 - \sigma_{se} N_2 \right) \right] P_s^{\pm} \mp \alpha P_s^{\pm} \quad (4)$$

Where α is the fiber loss; - and + is the backward and forward travelling of the signal power and the boundary conditions that govern them are as follow:

$$P_s^{+}(0) = R_i P_s^{-}(0) \quad (5)$$

$$P_s^{-}(L) = R_o P_s^{+}(L) \quad (6)$$

Here L denotes the length of the laser cavity; Fiber Brag Gartering reflectivity is denoted as R_i and R_o, where the former refers to for the input of the praseodymium-doped fiber and the later refers to the output. Equations (1)–(4) could be solved by applying the Relaxation algorithm approach. The signal power at the λ_s wavelength is propagated forth and back between the FBGs at the ends of the fiber. This process is iteratively repeated until the required accuracy meets the given conditions.

We apply the Relaxation approach to solve the pump differential equation eq.(6) and signal differential equation eq.(7). Table (2) summarizes all of the required parameters in the numerical investigation. It is assumed that the pump possesses loss levels and asignal wavelength to be 1 dB/m. This value exceeds the one adopted in the experimental investigations at the same host material [22].

Table 2. Values of numerical parameters used in the simulation set up [9, 12]

Quantity	Value	Unit
Lifetime of level 2	12	ms
Lifetime of level 3	4.2	ms
The concentration of Pr^{3+}-ion	5×10^{25}	m^{-3}
Branching ratio for the transitions between the level $^3H_6$ to $^3H_5$	0.42	
Signal (1) and Signal (2) wavelength	4.5	μm
Pump wavelength	2.04	μm
The field area of the effective mode at a certain pumping wavelength	812	μm^2
The field area of the effective mode at certain signal wavelength	979	μm^2
Confinement factor at the pump wavelength	0.73	
Confinement factor at the signal wavelength	0.62	
Absorption cross section at the pump wavelength	2.2×10^{-24}	m^2
The cross section of the emission at the pump wavelength	1.4×10^{-24}	m^2
The cross section of the emission at the signal (1) wavelength	0.15×10^{-24}	m^2
The cross section of the absorption at signal (1) wavelength	0.01×10^{-24}	m^2
The cross section of the emission at the signal (2) wavelength	0.66×10^{-24}	m^2
Absorption cross section at the signal (2) wavelength	0.75×10^{-24}	m^2

4. RESULTS

After optimizing the core fiber diameter and computing both the factor of confinement as well as the field area of the effective mode,the laser performance of the suggested fiber can be now investigated for MIR emission at 4.5µm wavelength. The rate equations of the laser model can be solved numerically by using the data of table (2). Then, several vital parameters are investigated to examine the resulted performance, these parameters include the length of the laser cavity and the pumping power for Pr^{3+}-doped chalcogenide laser based on W-type fiber. We assume that the proposed fiber laser is core pumped at one side while the MIR light is emitted from the other side. The input reflectivity is 0.05 and 0.95 for the pump and the signal, respectively whilst the reflectivity of the output FBG for the signal is 0.05. Firstly, the laser cavity length is investigated to find the optimized value. Fig. 6 displays the output laser power with the length of laser cavity at different values of the pump powers. It is clearly shown that the laser power increases in accordance with the increase of the cavity length for all the cases of the pump powers until reaching its maximum value, and then drops because of the fiber loss. The results also indicate that the required cavity length increases with the pump power and 0.7m is the optimized value at 1 W of the pump power and it is possible that more than 0.6 W of laser output power can be achieved at 4.5µm wavelength. The corresponding power intensity is about 1231MW/m^2, which is much less than the applied values in the other previous literature studies.

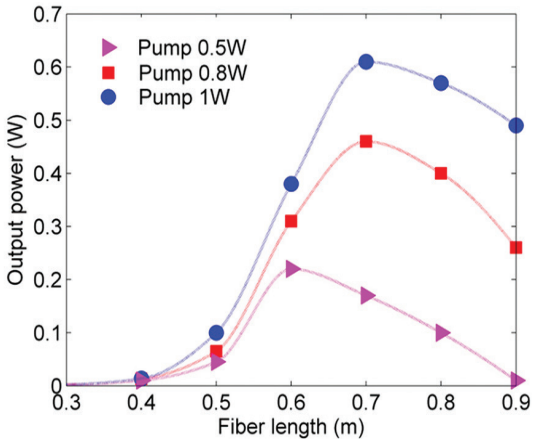


Fig. 6. The achievable output laser power as compared to the laser cavity-length evaluated atvarious

Next, the mechanism to obtain the efficiency of the laser slope is addressed, which is defined with the input pump power. In Fig. 7, an illustration of the dependency of the laser slope efficiency is depicted. This illustration is maintained upon the pumping power of two different fiber losses. The results indicate that more than 61% of slope efficiency is possible to achieve when the pump power is equal to 1W at fiber loss of 1dB/m. This slope drops to only 42% at the fiber loss of 2dB/m. Thus, it is possible by applying our proposed fiber to get laser emission at 4.5µm with high levels of fiber losses.

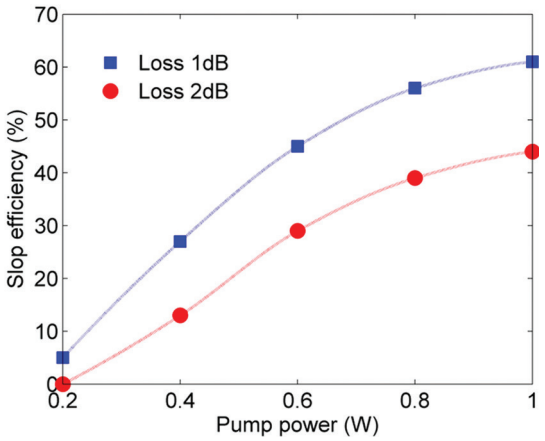


Fig. 7. The efficiency of the output slope against the applied pumping power measured at two values of fiber loss.

Another assessment of this paper is to evaluate the results of our design with the previously reported findings at MIR chalcogenide fiber lasers. We draw a comparison of W-type index pr^{3+}-doped chalcogenide fiber laser with previously published results as illustrated in table (3). The results of our design exhibit high laser efficiency than the first three reported MIR sources with lower applied pump intensity. In spite of the fact that the PCF design produces the highest slope efficiency, this design still exhibits the disadvantages of high complexity in the fabrication and the splicing. As a result, our proposed W-type fiber offers a good and robust way to release an efficient MIR coherent source with less complexity in fabrication, design and splicing which is more suitable for the laser systems.

Table 3. Comparing the obtained simulation outcome against reported results in existing literature.

Layout of the laser	Area (µm²)	Pumping intensity (MW/m²)	Efficiency of the slope
SIF with cascading laser concept [9]	2827	1768	16%
SIF with overlapped scheme [12]	2827	1768	48%
Multi-mode SIF based on resonant pumping design [15]	707	1414	54%
PCF with overlapping approach [16]	2050	500	64%
W- index type fiber with overlapping approach (current design)	812	1231	61%

5. CONCLUSIONS

In this study, a W- type index Pr^{3+}-doped chalcogenide fiber laser has been proposed for MIR laser source at 4.5µm wavelength. The proposed fiber exhibits the advantages of increasing the damage threshold of the pump power and restricting any nonlinearity effect by enlarging the effective mode field area. The proposed fiber consists of Pr^{3+}: GeAsGaSe core glass, GeAsGaSe-S glass in both the inner and the outer cladding but with different atomic level of Sulphur (S). The FEM is adopted to simulate the distribution of the optical basic mode of a W-typed fiber. The field area of the effective mode is about 840 μm^2 at the pump 2.04 µm wavelength and 940 μm^2 at MIR laser wavelength 4.5 µm.

In the laser layout, we avoid the need of cascaded FBGs for MIR fiber laser by exploiting the characteristic of the overlapping region in the emitting cross-sections that belongs to the chalcogenide fiber that is doped with Pr^{3+}. A single pair of FBGs is used at 4.5µm wavelength for both signal (1) and signal (2). This laser layout is not only avoiding the fabrication complexity of cascaded FBGs but also increases the slope efficiency by emitting two photons in MIR wavelengths from one excited ion. The simulation findings indicate that at 4.5µm wavelength, it could be seen that the realized slope efficiency exceeds more than 61% when the loss in the fiber is about 1dB/m. As a result, our proposed W-type fiber is robust and represent an effective way to release an efficient MIR coherent source with less complexity in fabrication, design and splicing which is more suitable for all laser systems.

6. REFERENCES

[1] V. A. Serebryakov, E. V. Boĭko, N. N. Petrishchev, A. V. Yan, "Medical applications of mid-IR lasers problems and prospects", Journal of Optical Technology, Vol. 77, 2010, pp. 6-17.

[2] A. B. Seddon, "Mid-infrared (IR) - A hot topic: The potential for using mid-IR light for non-invasive early detection of skin cancer in vivo", Physica Status Solidi B, Vol. 250, 2013, pp. 1020-1027.

[3] F. Starecki et al. "Mid-IR optical sensor for CO2 detection based on fluorescence absorbance of Dy3+: Ga5Ge20Sb10S65 fibers", Sensors and Actuators B: Chemical, Vol. 207, 2015, pp. 518-525.

[4] B. M. Walsh, H. R. Lee, N. P. Barnes, "Mid-infrared lasers for remote sensing applications", Journal of Luminescence, Vol. 169, 2016, pp. 400-405.

[5] D. Caffey et al. "Recent results from broadly tunable external cavity quantum cascade lasers", Proceedings Volume 7953, Novel In-Plane Semiconductor Lasers X, 16 February 2011, p. 79531K.

[6] M. Ebrahimzadeh, "Mid-infrared ultrafast and continuous-wave optical parametric oscillators", Solid-State Mid-Infrared Laser Sources, Springer, Berlin, Heidelberg, 2003, p.179.

[7] A. A. Kosterev et al. "Methane concentration and isotopic composition measurements with a mid-infrared quantum-cascade laser", Optics Letters, Vol. 24, 1999, pp. 1762-1764.

[8] M. A. Khamis, R. Sevilla, K. Ennser, "Design of W-Type Index Chalcogenide Fiber for Highly Coherent Mid-IR Supercontinuum Generation", Journal of Lightwave Technology, Vol. 36, 2018, pp. 5388-5394.

[9] Ł. Sójka et al. "Study of mid-infrared laser action in chalcogenide rare earth doped glass with Dy3+, and Tb3+", Optical Materials Express, Vol. 2, 2012, pp. 1632-1640.

[10] M. C. Falconi et al. "Design of an Efficient Pumping Scheme for Mid-IR Dy3+:Ga5Ge20Sb10S65 PCF Fiber Laser", IEEE Photonics Technology Letters, Vol. 28, 2016, pp. 1984-1987.

[11] S. D. Jackson, T. A. King, M. Pollnau, "Diode-pumped 1.7-W erbium 3-µm fiber laser", Optics Letters, Vol. 24, 1999, pp. 1133-1135.

[12] M. A. Khamis, K. Ennser, "Design of Highly Efficient pr+3-doped Chalcogenide Fiber Laser", IEEE Photonics Technology Letters, Vol. 29, 2017, pp. 1580-1583.

[13] L. D. Merkle, Z. Fleischman, E. E. Brown, J. L. Allen, U. Hommerich, M. Dubinskii, "Enhanced mid-infrared emission of ions in solids through a "3-for-1" excitation process – quantified", Optics Express, Vol. 29, 2021, pp. 39001-39015.

[14] L. B. Shaw, B. Cole, P. A. Thielen, J. S. Sanghera, I. D. Aggarwal, "Mid-wave IR and long-wave IR laser potential of rare-earth doped chalcogenide glass fiber", IEEE Journal of Quantum Electronics, Vol. 37, 2001, pp. 1127-1137.

[15] L. Sójka et al. "Numerical and experimental investigation of mid-infrared laser action in resonantly pumped doped chalcogenide fibre", Optical and Quantum Electronics, Vol. 49, 2017, p. 11426.

[16] M. A. Khamis, R. Sevilla, K. Ennser, "Large Mode Area Pr+3-Doped Chalcogenide PCF Design for High Ef-

ficiency Mid-IR Laser", IEEE Photonics Technology Letters, Vol. 30, 2018, pp. 825-828.

[17] J. Kim, C. Codemard, Y. Jeong, J. Nilsson, J. K. Sahu, "High power continuous-wave Yb-doped fiber laser with true single-mode output using W-type structure", Proceedings of the Conference on Lasers and Electro-Optics/Quantum Electronics and Laser Science Conference and Photonic Applications Systems Technologies, Technical Digest, Long Beach, CA, USA, 21-26 May 2006.

[18] M. A. Arbore, Y. Zhou, H. Thiele, J. Bromage, L. Nelson, "S-band erbium doped fiber amplifiers for WDM transmission between 1488 and 1508 nm", Proceedings of the OFC 2003 Optical Fiber Communications Conference, Atlanta, GA, USA, 28 March 2003.

[19] S. Yoo et al. "Analysis of W-type waveguide for Nd-doped fiber laser operating near 940 nm", Optics Communications, Vol. 247, 2005, pp. 153-162.

[20] A. S. Vakhrushev et al. "W-type and Graded-index bismuth-doped fibers for efficient lasers and amplifiers operating in E-band", Optics Express, Vol. 30, 2022, pp. 1490-1498.

[21] Z. Tang et al. "Mid-infrared photoluminescence in small-core fiber of praseodymium-ion doped selenide-based chalcogenide glass", Optical Materials Express, Vol. 5, 2015, pp. 870-886.

[22] G. E. Snopatin, V. S. Shiryaev, V. G. Plotnichenko, E. M. Dianov, M. F. Churbanov, "High-purity chalcogenide glasses for fiber optics", Inorganic Materials, Vol. 45, 2009, pp. 1439-1460.

Harmonic reduction using Particle Swarm Optimization based SHE Modulation Technique in Asymmetrical DC-AC Converter

Original Scientific Paper

Kamaldeep Boora
Assistant Professor,
School of Electrical Engineering,
Shri Mata Vaishno Devi University, Katra-182320
kamaldeep@smvdu.ac.in

Isha Malhotra
Assistant Professor,
School of Languages and Literature,
Shri Mata Vaishno Devi University, Katra-182320
isha.malhotra@smvdu.ac.in

Anurag Kumar
Assistant Professor,
School of Languages and Literature,
Shri Mata Vaishno Devi University, Katra-182320
anurag.kumar@smvdu.ac.in

Vikram Kumar
Researcher,
School of Electrical Engineering,
Shri Mata Vaishno Devi University, Katra-182320
vikram.kumar@smvdu.ac.in

Abstract *– Many inverters play an important role in transmitting and processing energy to power system networks. To reduce the cost and size of multilevel inverters, various topologies have been included in the literature. But these topologies do not look at the complete harmonic distortion in the output waveform. In this study, a modern multilevel inverter structure, a mutated H-bridge inverter is adopted that demands a small amount of switches, driver circuits, power diodes and DC voltage sources compared to conventional multilevel inverters to produce the required level in the output voltage. The mutated H-bridge converter uses a nearest-level control method, produces high value of total harmonic at the output voltage and low-level harmonics content is also high, which is more dangerous than high-order harmonics. Therefore, the selective harmonic elimination (SHE) method is used to reduce the low frequency harmonics and the total harmonic distortion to the output voltage. Comparison of complete harmonic deviation and low-level harmonic content using the above-mentioned control strategies on the 31-level inverter is presented. Simulation studies confirm the performance of a 31-level inverter with low-order harmonics and a complete harmonic distortion using the SHE process. The effectiveness and accuracy of the SHE in producing 31 level waveform is demonstrated by utilizing the test outcomes and the THD found is of the order of 2.8%, 1.7%, 1.7% and 1% for the modulation index values of 0.5590, 0.7440, and 0.8380 and 0.9110 respectively.*

Keywords*: CHB, Selective harmonic elimination technique, Asymmetric multilevel inverters, Nearest level control technique.*

1. INTRODUCTION

Multilevel converters for power conversion were first introduced in the last 30 years [1]. A common theory involves the use of a large number of DC power sources and power switches to perform power conversion. In multi-level inverters, as the number of DC power sources and power switches increases, the output waveform voltage approaches almost as a sinusoidal. Compared to a two-level inverter, the main advantages of multilevel inverters are low switching losses [2], low electrical stress on every device [3], better electronic magnetic compatibility [4] and low harmonic content in the output wave [5]. The main advantage of a multilevel inverter over traditional inverters is that it can operate at high voltage using low-voltage semiconductor switches [6], [7].

Most inverters can be divided into two categories: conventional DC sources and cascaded inverter with separate DC sources. Diode clamped inverter and flying capacitor inverter fall into the conventional category of DC sources 1], [8]. In the diode clamped inverter topology, the number of capacitors in the series is used to separate the DC bus voltage according to the step required for the output voltage waveform. But as the levels in the output voltage increases, the voltage sharing between the different capacitors becomes uneven. To overcome this problem, clamping diodes are used, but due to the use of these diodes, the cost and size of the inverter increases. In flying capacitor topology, capacitors are used as a binding device. In this topology, the voltage at every capacitor is different from that of the next capacitor. The magnitude of the volt-

age level in the output waveform depends on the differences between the voltages of the two capacitors. To produce m-levels of output voltage, m-1 capacitors are required in DC bus.

The Cascading H-bridge (CHB) inverter belongs to the second category [9], [10]. In this topology, various H-bridges with a separate DC source are connected in cascading to increase the waveform steps. This topology requires the same amount of electronic power switch as the diode clamped topology, but does not require clamping diodes. The main advantages of this topology is modularity in structure, simple control method and number of output levels can be increased or decreased by simply increasing and removing the H-bridge cell. But the formation of CHB requires a large number of power converters, DC power sources and driver circuits to produce a high level of output voltage.

Many topology having reduced number of power switches, DC power sources and driver circuits developed in [11] - [15]. In [11], asymmetric topology is presented, but the main disadvantage of this topology is the use of dual-directed switches, which increase the inverter cost. The three topology of multilevel inverter introduced in [12], are supported as diode clamped, symmetrical and asymmetrical cascaded multilevel inverter (CMLI). However the use of a transformer and rectifier on the input side to achieve a high DC voltage makes its size larger. In [13], two different algorithms are introduced in symmetric and asymmetric multilevel inverter topology. However, a high number of IGBTs were used as this topology uses dual-directed switches which are the main features of this structure.

In this study, a modern mutated H-bridge [16] is adopted to obtain an increased count of output voltage steps with a shortened number of power switch devices compared to traditional CMLI. This structure, which uses a nearest-level control (NLC) system, produces an output wave with high-magnitude harmonic content. Therefore, the selective harmonic elimination (SHE) control method is adopted to reduce or lower the harmonic content and the total harmonic distortion (THD) in the output voltage waveform. Solution for non linear equation of SHE technique describing the harmonic contents, can be solved by Resultant theory [17], Genetic Algorithm [18], theory of symmetric polynomials [19], Newton-Raphson method [20], or other appropriate means. However, as the number of levels increases, the number of polynomials in mathematics becomes so large that solving them becomes more difficult and the quality of the solution deteriorates. Therefore, the Particle swarm optimization development process is adopted in this study to calculate the shooting angle of the SHE method as it has the surface exploring power at the commencing of the run and the local exploring near the end of the run. The comparative effect of THD on the output voltage using both methods has been introduced to investigate the height of SHE efficiency. Lastly, the realization of the mutated H-bridge inverter to produce all levels of voltage using the SHE method is verified using MATLAB / SIMULINK and the experimental result of the 31-level inverter structure.

2. MUTATED H-BRIDGE INVERTER

The elemental unit in this topology is a combination of two switches and one DC voltage source with traditional H-bridge cell or it can be said that it is the combination of two DC voltage source (V_1 and $V_{1'}$) and six unidirectional power electronic switches (S_1, S_2, $S_{1'}$, $S_{2'}$, S_a and S_b) . Fig. 1 shows the elemental unit for the mutated H-bridge inverter which is able to generate 7-level in output waveform [16].

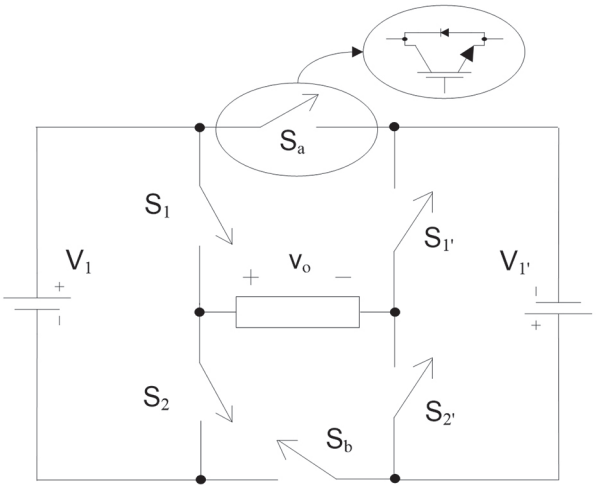


Fig. 1. Mutated H-bridge inverter structure (elemental unit) [16]

To achieve all the levels in the output waveform, the power sources V_1 and $V_{1'}$, must be in the ratio 2:1 or V_1 must be equal to the $2V_{1'}$.

The 31-level inverter structure can be upgraded using an elemental unit by connecting additional power switches and DC power sources in the manner shown in Fig. 2. This contains four DC power sources and ten different power switches. According to Fig. 2, synchronous switching (S_1, S_2), (S_3, S_4), ($S_{1'}$, $S_{2'}$) and ($S_{3'}$, $S_{4'}$) lead to shorting of DC power sources V_1, V_2, $V_{1'}$ and $V_{2'}$. Therefore, synchronous opening of these transitions must be averted. Concurrently, synchronous opening of S_a and S_a should be averted. The switch is operated as per [16] to obtain the output waveform of the multilevel inverter of the 31-level inverter.

The size of the DC voltage supplies for configuration as depicted in Fig. 2 are related by [16]:

$$V_{1'} = V_{DC} \quad (1)$$

$$V_1 = 2V_{DC} \quad (2)$$

$$V_{2'} = 5V_{DC} \quad (3)$$

$$V_2 = 10V_{DC} \quad (4)$$

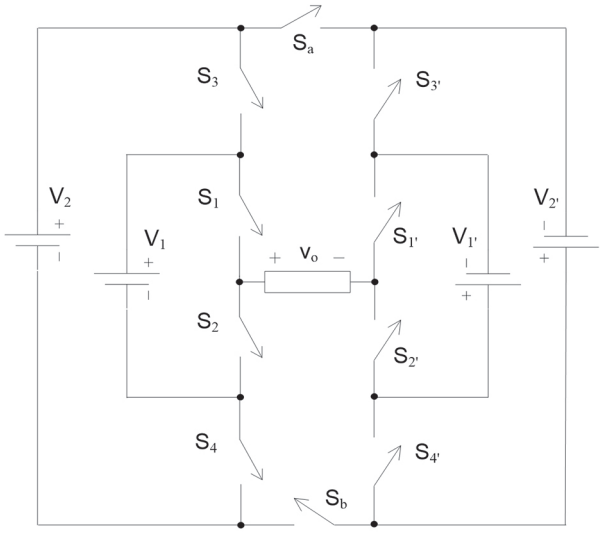


Fig. 2. 31-level inverter structure

3. MODULATION TECHNIQUES

In the literature, many variables have been introduced and can be categorized as low frequency method and high frequency method [6]. In low frequency switching methods, there is very little switching of equipment compared to high frequency methods where switching equipment is changed multiple times. The NLC, SHE, space vector control etc. comes with a low frequency change while the phase shifted, phase disposition, phase opposition disposition, space vector PWM etc. falls into the path of the most changing frequency. As the switches change frequently, switching losses increase in a highly variable mode. Therefore, the NLC and SHE strategies are adopted in this study to swap switches in a 31-level inverter structure.

3.1. NEAREST LEVEL CONTROL (NLC) TECHNIQUE

In this process [4], the reference switch signal is compared to the available voltage that can be generated and switches corresponding to that level are turned on. Here, v^* is the reference signal and the V_{aN} is the nearest available level of voltage. The output voltage can be calculated using the following equation:

$$v=V_{DC}\times round\{v^*/V_{DC}\} \quad (5)$$

Where v is the output voltage level, v^* is the reference voltage and V_{DC} is the difference between two consecutive voltage level.

A circular function is used to round the value of a variable to the nearest number [4]. Suppose, if the value of a variable is 10.7 then round it to the next value as 11 and if the value of the variable is 10.4 then round it up to the next minimum value as 10. Further details on this process can be found at [4].

3.2. SELECTIVE HARMONIC ELIMINATION TECHNIQUE

In this approach, the switching angles are predetermined in a manner to eradicate the lower magnitude of harmonics and to achieve a basic fraction equal to the reference voltage [4], [6], [21]. As an example of a seven-level inverter, the output waveform is characterized in mathematical form by extending the Fourier Series as [4], [21], [23]:

$$v_{an}(\omega t) = \sum_{k=1,3,5\ldots}^{\infty}\left(\frac{4V_{dc}}{k\pi}\right)(cos(k\alpha_1)+cos(k\alpha_2) + cos(k\alpha_3))\,sin(k\omega t) \quad (6)$$

where V_{DC} is the size of the independent DC voltage source, α_1, α_2 and α_3 are the switching angles for 1st, 2nd and 3rd unit. Switching angles to diminish the harmonics of the lower order should be calculated in such a way that $0 \le \alpha_1 < \alpha_2 < \alpha_3 \le \pi/2$. With the s number of angles, the same number of degrees of freedom is obtained, one of which is used to obtain the required output voltage and the remaining free degrees (s-1) can be used to eliminate sub-system harmonics. Using (6), the basic output voltage can be recorded by relating the switching angles as follows (assuming $k = 1$):

$$(4V_{DC}/\pi)(cos\,(\alpha_1)+cos(\alpha_2)+cos(\alpha_3)) = V_f \quad (7)$$

Magnitude of basic voltage is associated with utmost attainable output voltage by a term named modulation index m_I and can be denoted as:

$$\mathrm{m_I} = \pi \mathrm{V_f}/(4\mathrm{sV_{DC}}) \quad (8)$$

Where s is the total count of angles, V_{DC} is the DC voltage supply for every unit and V_f is the basic output voltage.

As even harmonics are absent due to the half wave symmetry of the waveform and triplen harmonics are vanished in line to line voltage. In general, the non-triplen odd harmonic components need to be eliminated or minimized. The subsequential mathematical statements are adopted to determine the switching angles:

$$\begin{aligned} \cos\,(\alpha_1)+\cos(\alpha_2)+\cos(\alpha_3) &= 3\mathrm{m}_1 \\ \cos\,(5\alpha_1)+\cos(5\alpha_2)+\cos(5\alpha_3) &= 0 \\ cos\,(7\alpha_1)+cos(7\alpha_2)+cos(7\alpha_3) &= 0 \end{aligned} \quad (9)$$

With the given value of m_I (0 to 1), value of α_1, α_2 and α_3 are calculated by diminishing the 5th and 7th harmonic components with minimal complexity and load calculation.

To define the degree of waveform that is deviated from the appropriate sine wave, using a reference term also named total harmonic distortion (THD). It is described as the square root measure of sum of square of high frequency element to that of the basic frequency element [4], [21]:

$$THD_{49} = \frac{\sqrt{V_5^2+V_7^2+\cdots+V_{49}^2}}{V_1}\times 100 \quad (10)$$

where V_1, V_5, V_7,…,V_{49} are the magnitude of the amplitude of basic, 5th, 7th, …, 49th harmonic components

respectively. The mathematical equations as expressed in (9) are solved using the Particle Swarm Optimization (PSO) method [22].

PSO is a stochastic community conduct based on the particular and group conduct of different breed in the atmosphere in seeking of food or their requirement [22]. In this process, the particles sail in different direction in atmosphere and shift their location agreeing to their familiarity and agreeing to the surrounding particles, thus using the finest environment in which they interact with themselves and their neighbors. Every character has its current location X_i=[x_{i1} , x_{i2},…x_{iD}] and a velocity V_i= [v_{i1} ,v_{i2} ,…v_{iD}], where D is the size of the search space. Particle speed indicates a change in location from one step to another. Every character remembers its personal finest location (p_{best}) according to the finest qualification value in the search field. Global finest value (g_{best}) is the finest value obtained by one of the characters in a group. During testing, characters adjust their position and speed from their experience and previous knowledge of other characters in the group. This means that each character changes its location and speed depending on the span between its present location and pbest and the span between its present location and g_{best}. The finest location detected by ith character is recorded and represented as $P_{best,i}$=[$P_{best,i1}$, $P_{best,i2}$, …$P_{best,iD}$]. The finest location found by another character among group characters represented as G_{best} =[G_{best1} , G_{best2},…G_{bestD}] Updated speed and status of i^{th} character in d^{th} dimension using present speed and distance from $P_{best,iD}$ and $G_{best,D}$ can be given as:

$$\mathrm{v_{iD}(t+1)} = \chi[\mathrm{v_{iD}(t)} + \mathrm{c_1 r_1}\left(\mathrm{P_{best,iD}} - \mathrm{x_{iD}(t)}\right) + \mathrm{c_2 r_2}\left(\mathrm{G_{best,D}} - \mathrm{x_{iD}(t)}\right)] \quad (11)$$

$$\mathrm{x_{iD}(t+1)} = \mathrm{x_{iD}(t)} + \mathrm{v_{iD}(t+1)}\ \mathrm{i} = 1,2,\ldots \mathrm{m} \quad (12)$$

where m is the sum of characters in group, t is the sum of repetition (generation), D is the sum of dimensions corresponds to sum of members of every character, v_{iD} (t+1) is the velocity of member D of character i at repetition t+1, v_{iD} (t) is the velocity of member D of character i at repetition t, x_{iD} (t+1) is the position of member D of character i at repetition t, where c_1 and $c2$ are the cogitative and group parameters respectively, $r1$ and $r2$ are the arbitrary numbers between 0 and 1, $P_{best,iD}$ is the local finest location of member D of character i.

A new factor named constriction factor χ has been inserted to enhance the convergence attribute of PSO algorithm as:

$$\chi = \frac{2\mathrm{k}}{\left|2 - \mathrm{c} - \sqrt{\mathrm{c}^2 - 4\mathrm{c}}\right|}, \mathrm{c} = \mathrm{c_1} + \mathrm{c_2}; \mathrm{c} > 4. \quad (13)$$

The calculated switching angles should produce desired basic output voltage and be in the range of 0 to $\pi/2$. In this work, the cost function used to reduce the harmonic content in the output voltage is given as follows:

$$\mathrm{f}(\alpha_1, \alpha_2, \alpha_3, \ldots, \alpha_{15}) = \mathrm{abs}\left(\mathrm{M} - \mathrm{abs}(\mathrm{v_1})\right) + \mathrm{abs}(\mathrm{v_5}) + \mathrm{abs}(\mathrm{v_7}) + \cdots + \mathrm{abs}(\mathrm{v_{43}}) \quad (14)$$

Where M=m_I * $4sV_{DC}/\pi$ and α_1, α_2, α_3,…, α_{15} are the switching angles.

4. SIMULATION RESULTS

In this section, the performance of the above-mentioned topology is validated by simulation studies employing NLC and SHE techniques for 31-level inverter configuration as depicted in Fig. 2. For NLC technique, the utmost amplitude of reference signal plays an important role in deciding the magnitude of output basic voltage However, as the magnitude of the reference signal changes, the THD at the output voltage also changes. The effectiveness of the NLC strategy is confirmed by conducting simulation studies using the parameters V_1, V_2, $V_{1'}$ and $V_{2'}$ are 30V, 150V, 15V and 75V with R–L load having R = 45Ω and L = 55mH (as stated in [16]). The output phase voltage and the current waveform obtained using the NLC methods are shown in Fig. 3. Fourier analysis of the waveforms shown in Fig. 3 is given in Fig. 4.

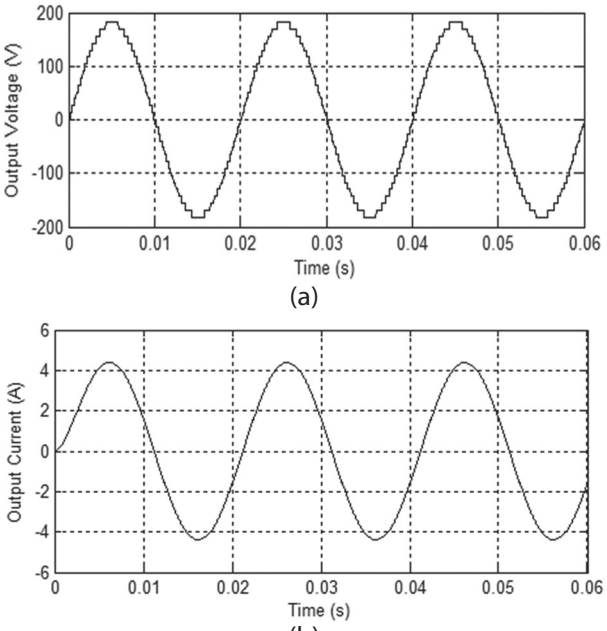
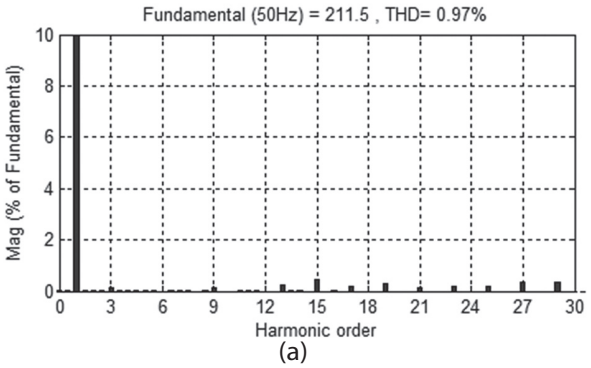
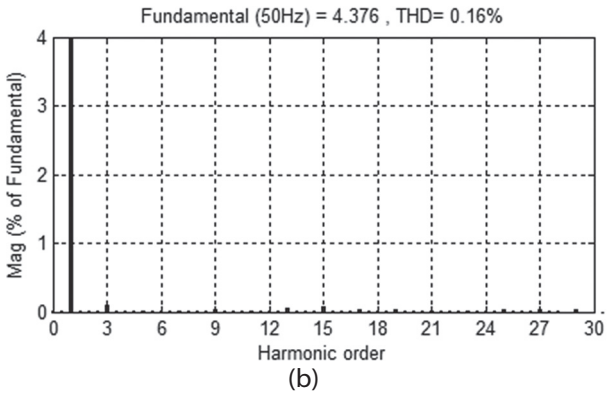


Fig. 3. Simulated waveform using NLC technique for (a) phase voltage (b) phase current.





(b)

Fig. 4. THD in output (a) phase voltage (b) phase current.

It is observed that THD values in the generated phase and current are still 0.97% and 0.16% respectively and these values are comparable to the results given in [16], i.e. 0.94% and 0.19% respectively. Therefore, the effectiveness of the NLC strategy adopted in this project is validated by simulation studies.

Using the SHE modulation scheme, the angle conversion values were obtained using the PSO method for the various modulation index values as shown in Table 1 of the 31-level inverter. For real-time use of the SHE system, changing the angles corresponding to the modulation index can be kept in the look-up table and can be changed easily and efficiently.

To ensure computerized angles using the PSO method, simulation courses in the MATLAB / SIMULINK field were developed for a 31-level converter with the following parameters: elemental frequency = 50Hz, V_1, V_2, $V_{1'}$ and $V_{2'}$ are 24V, 120V, 12V and 60V respectively (statistically (1) - (4)).

A correlation study of calculation and simulation outcomes is presented in Table 2. As shown in Table 2, for example, in m_I= 0.7440, the magnitude of the 5th and 7th harmonic components by calculation were 0.03% and 0.39% respectively while in simulation these numbers are 0.02% and 0.42% respectively. The results obtained with numeracy and simulation studies are very close. Concurrently, in some m_I values the results are very consistent. Therefore, these outcomes validate the calculation results obtained by the PSO method using simulation studies.

In order to ensure the superiority of the SHE strategy over the NLC method, simulation studies of different modulation index values were performed using the NLC method. Fig. 5 (a) shows the line voltage waveform generated using the NLC method while the corresponding waveform using the SHE method is shown in Fig. 5 (b) for m_I= 0.7440. It can be seen from these statistics that the desired levels are obtained using both methods.

Fourier analysis of line voltage obtained by NLC and SHE strategies are shown in Fig. 6 (a) and Fig. 6 (b) respectively. It can be seen from these statistics that THD values are 1.21% and 0.83% respectively in NLC and SHE's for basic output voltage of 229V.

In addition, comparative studies with respect to the lower order of harmonics shows that the magnitude of the 5th and 7th part of the waveform obtained using NLC is 0.59% and 0.43% respectively, whereas in the case of SHE these values are 0.02% and 0.42%, very small compared to the NLC method.

Table 1. Swapping angles for 31-level inverter employing SHE

Swapping angles	Modulation Index			
	0.5590	0.7440	0.8380	0.9110
α_1	25.89	2.17	0	0.26
α_2	32.68	5.46	3.27	2.86
α_3	37.96	11.90	6.43	5.86
α_4	40.45	13.79	7.07	7.62
α_5	44.89	17.87	12.86	9.93
α_6	46.15	21.81	13.27	14.25
α_7	50.85	25.76	18.37	14.57
α_8	52.89	29.50	22.86	20.10
α_9	57.10	35.51	26.82	20.92
α_{10}	61.15	41.17	31.75	26.26
α_{11}	64.38	48.86	35.48	26.28
α_{12}	67.27	53.29	39.34	32.74
α_{13}	71.99	59.75	42.33	37.53
α_{14}	76.81	67.60	50.29	41.31
α_{15}	81.22	89.98	78.96	49.05

Table 2. Correlation of calculated (C) and simulated (S) outcomes

Modulation Index	0.5590		0.7440		0.8380		0.9110	
	C	S	C	S	C	S	C	S
Fundamental r.m.s voltage (V)	171.5	171.5	226.3	229	256.4	257.4	276.2	275.9
5th Harmonic content (%)	0.05	0.19	0.03	0.02	0.07	0.08	0.13	0.26
7th Harmonic content (%)	0.06	0.15	0.39	0.42	0.11	0.20	0.16	0.12
11th Harmonic content (%)	0.03	0.05	0.04	0.11	0.03	0.04	0.01	0.09
Total Harmonic Distortion (%)	0.95	1.01	0.79	0.83	0.80	0.91	0.84	0.88

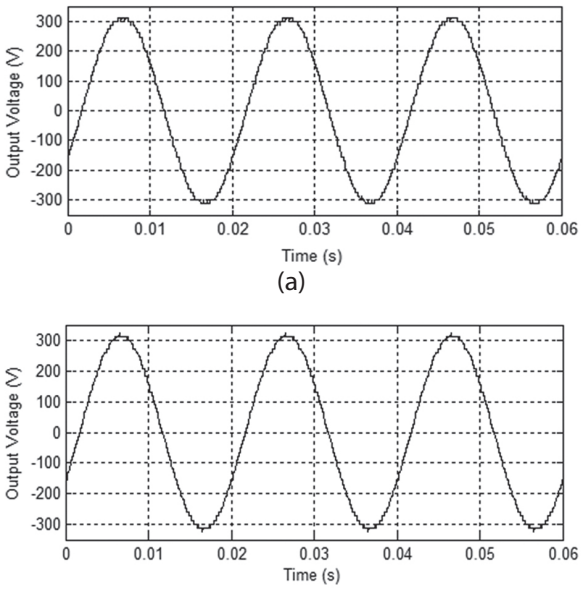


Fig. 5. Simulated waveform of output line voltage at $m_I = 0.7440$ employing (a) NLC (b) SHE.

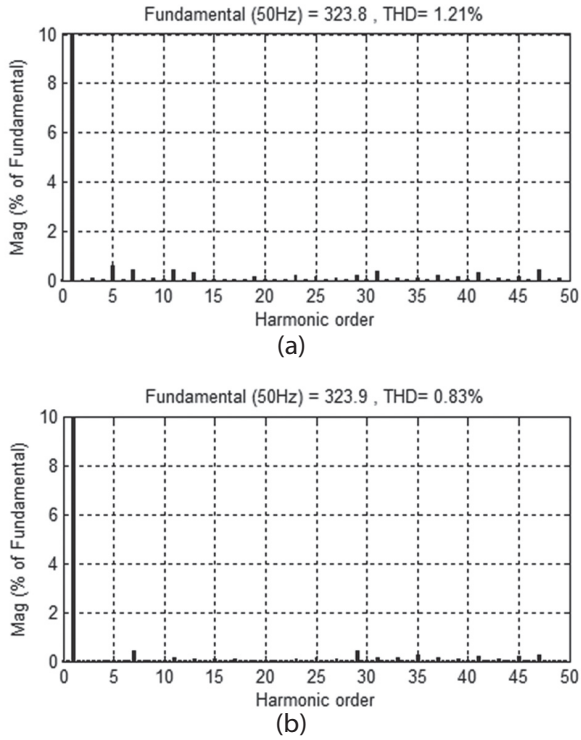


Fig. 6. Harmonic spectrum of output line voltage at m_I 0.7440 employing (a) NLC (b) SHE.

As shown in Table 3, in m_I= 0.8380, the magnitude of the 5th and 7th harmonic components using NLC is 2.63% and 1.21% respectively while using the SHE process module, 5th and 7th harmonics reduced to 0.08% and 0.20% respectively. The comparative data between the two methods is shown in Table 3.

Therefore, the above study shows that THD and the magnitude of the lower order harmonics in line voltage are much smaller in the SHE as compared to NLC. It may be noted that in the harmonic spectrum the utmost values of the basic voltage are displayed.

Table 3. Comparison of THD in output line voltage employing NLC and SHE technique

Modulation Index	0.5590		0.7440		0.8380		0.9110	
	NLC	SHE	NLC	SHE	NLC	SHE	NLC	SHE
Fundamental r.m.s voltage (V)	172.4	171.5	229	229	258.3	257.4	276.5	275.9
5th Harmonic content (%)	0.53	0.19	0.59	0.02	2.63	0.08	2.50	0.26
7th Harmonic content (%)	0.45	0.15	0.43	0.42	1.21	0.20	1.94	0.12
11th Harmonic content (%)	0.43	0.05	0.42	0.11	0.53	0.04	0.81	0.09
Total Harmonic Distortion (%)	2.07	1.01	1.21	0.83	3.03	0.91	3.33	0.88

Furthermore, in order to find out distortion in output current, a performance parameter i.e. weighted total harmonic distortion (WTHD) is calculated and defined as follow [23]:

$$WTHD_{49} = \frac{\sqrt{(V_5/5)^2+(V_7/7)^2+\cdots+(V_{49}/49)^2}}{V_1} \times 100 \quad (15)$$

Table 4. Comparison of WTHD in output line voltage

Modulation Index	NLC	SHE
0.5590	0.1522	0.0774
0.7440	0.0850	0.0588
0.8380	0.5726	0.0570
0.9110	0.5769	0.0635

The table 4 shows the comparison of WTHD in output line voltage employing the NLC and SHE technique. WTHD in case of SHE is very less as compare to NLC and comparative result in Table 4 validates the performance of multilevel inverter employing SHE technique.

5. EXPERIMENTAL RESULTS

A prototype setup is advanced in an effort to verify the diverse results received from the simulation studies as shown in Fig. 7 for 31-level inverter at m_I=0.7440. The details of numerous thing adopted for hardware implementation are as: Insulated Gate Bipolar Transistor (IGBT, FGW40N120HD) with a 1200V/40A capacity, DC voltage sources value V_1, V_2, $V_{1'}$ and $V_{2'}$ are 24V, 120V, 12V and 60V respectively. The switching pulses had been issued to

switching gadgets through dSPACE1104 and real time interfacing card. The basic output voltage of inverter may be managed in actual time via changing values of switching angles. A dead band is required to avert any short circuit due to synchronous flip-on of transfer pair (S_1, S_2) or (S_1, S_2) etc. A value of dead band relies upon the ON and OFF delay time of the electronic switch, minimum and maximum propagation delay of driver and a constant called safety margin. Considering these entire factor, a dead band of 5µsec is furnished to the switches pair. Four isolated channel digital storage oscilloscope, TPS2024B via TEKTRONIX is used for visualization of output voltage and Fluke 43B 600V/1250A is used for THD analysis in output voltage.

Fig. 7. Laboratary prototype.

A comparative study for output line voltage of 31-level inverter is carried out and has been provided in this segment. Evaluation of THD and lower order harmonic element is finished for non triplen harmonic components. Firing angles similar to numerous modulation indexes are used to operate the numerous switches. Output line voltage for 31-level inverter is shown in Fig. 8 and it could be seen that all positive and negative steps are present in the output waveform. Basic output line voltage and its THD at m_I 0.7440, 0.8380 and 0.9110 is shown in Fig. 9 (a)-(c) respectively. From Fig. 9 (a)-(c) it can be located that THD at m_I 0.7440, 0.8380 and 0.9110 is 1.7%, 1.7% and 1% respectively. A comparative result of simulation and experimental outcomes are presented in Table 5.

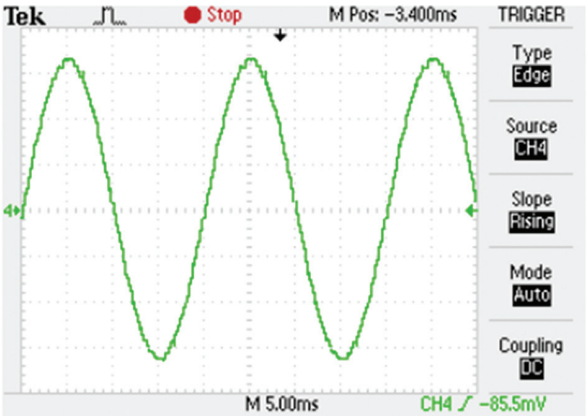


Fig. 8. Output line voltage waveform for 31-level inverter employing SHE technique at m_I= 0.7440.

A little distinction between the simulated and hardware outcomes is because of the different manner and platform adopted for these studies. Also, simulation and experimental outcomes of SHE are in near settlement and subsequently the validity of this topology with SHE technique is proved.

Table 5. Comparison of simulated (S) and experimental (E) results

Modulation Index	0.5590		0.7440		0.8380		0.9110	
	S	E	S	E	S	E	S	E
Fundamental r.m.s voltage (V)	171.5	170.3	229	225.7	257.4	254.9	275.9	276.2
5th Harmonic content (%)	0.19	0.50	0.02	0.30	0.08	0.20	0.26	0.10
7th Harmonic content (%)	0.15	0.50	0.42	0.50	0.20	0.30	0.12	0.10
11th Harmonic content (%)	0.05	0.50	0.11	0.40	0.04	0.30	0.09	0.10
Total Harmonic Distortion (%)	1.01	2.80	0.83	1.70	0.91	1.70	0.88	1.00

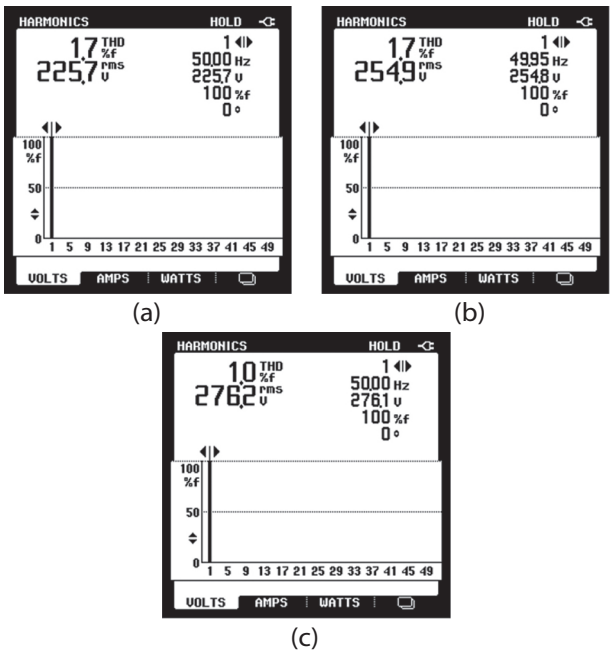
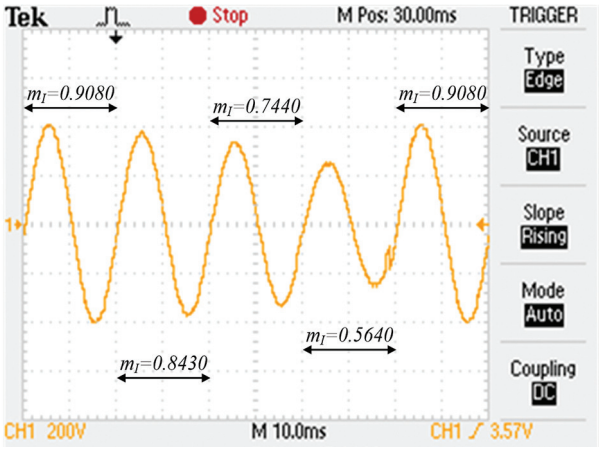


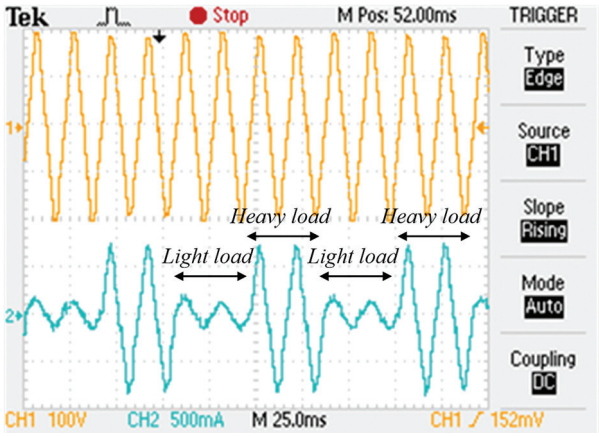
Fig. 9. 31-level output line voltage THD at m_I (a) 0.7440 (b) 0.8380 (c) 0.9110.

To analyze the dynamic behavior of this topology using the SHE method, a step change in the modulation

index and load variability is considered. The variation of the output voltage during the step change is shown in Fig. 10 (a). It is evident from this figure that there is a change in the magnitude of the output voltage (line voltage) as the m_I varies (first reduced and then increased). Load voltage and current waveforms during the change of load from light load to heavy load and from heavy load to light load are shown in Fig. 10 (b), and indicates that there is not much distortion in the output voltage waveform under the temporary condition.



(a)



(b)

Fig. 10. Output voltage waveform with (a) variation in modulation index (b) load variation.

6. CONCLUSION

In this study, a unique approach the usage of SHE method has been provided to limit the frequency component of low magnitude and total harmonic distortion in mutated H-bridge configuration. Particle swarm optimization based SHE technique is evolved for 31-level inverter. A comparative result using NLC and SHE method have been provided for 31-level inverter. The simulation and experimental results depict that proposed SHE method can diminish the lower order harmonics and THD in output voltage in comparison to NLC method. To verify the simulated consequences, experimental results are also provided. The THD found is of the order of 2.8%, 1.7%, 1.7% and 1% for the modulation index values of 0.5590, 0.7440, and 0.8380 and 0.9110 respectively using SHE for producing 31 level waveform as demonstrated by utilizing the test outcomes.This research is prospective to future studies within the subject of power conversion with high satisfactory output voltage as in renewable power where separate DC sources may be acquired by use of solar photovoltaic panel, fuel cell etc.

7. REFERENCES

[1] A. Nabae, I. Takahashi, H. Akagi, "A New Neutral-Point-Clamped PWM Inverter", IEEE Transactions on Industry Applications, Vol. IA-17, No. 5, 1981, pp. 518-523.

[2] J. W. Dixon, M. Ortuzar, L. Moran, "Drive system for traction applications using 81 level converter", Proceedings of the IEEE Vehicular Power Propulsion Conference, Paris, France, 6-8 October 2004.

[3] D. Zhong, L. M. Tolbert, J. N. Chiasson, B. Ozpineci, L. Hui, A. Q. Huang, "Hybrid cascaded H bridges multilevel motor drive control for electric vehicles", Proceedings of the 37th IEEE Power Electronics Specialists Conference, Jeju, Korea, 18-22 June 2006.

[4] J. Rodriguez, L.G. Franquelo, S. Kouro, J.I. leon, R.C. Portillo, M. A. M. Prats, M. A. Perez, "Multilevel converters: An enabling technology for high-power applications", IEEE Proceedings, Vol. 97, No. 11, 2009, pp. 1787-1817.

[5] Y. Tian, H. R. Wickramasinghe, P. Sun, Z. Li, J. Pou, G. Konstantinou, "Assessment of Low-Loss Configurations for Efficiency Improvement in Hybrid Modular Multilevel Converters", IEEE Access, Vol. 9, 2021, pp. 158155-158166.

[6] J. Rodriguez, J S Lai, F. Z. Peng, "Multilevel Inverters: A Survey of Topologies, Controls, and Applications", IEEE Transactions on Industrial Electronics, Vol. 49, No. 4, 2002, pp. 724-738.

[7] A. Salem, H. Van Khang, I. N. Jiya, K. G. Robbersmyr, "Hybrid Three-Phase Transformer-Based Multilevel Inverter with Reduced Component Count", IEEE Access, Vol. 10, 2022, pp. 47754-47763.

[8] R. Goel, T. T. Davis, A. Dey, "Thirteen-Level Multilevel Inverter Structure Having Single DC Source and Reduced Device Count", IEEE Transactions on Industry Applications, Vol. 58, No. 4, 2022, pp. 4932-4942.

[9] G. Schettino, C. Nevoloso, R. Miceli, A. O. D. Tommaso, F. Viola, "Impact Evaluation of Innovative Selective Harmonic Mitigation Algorithm for Cascaded H-Bridge Inverter on IPMSM Drive Application", IEEE Open Journal of Industry Applications, Vol. 2, 2021, pp. 347-365.

[10] J. Kumar, P. Agarwal, B. Das, "Implementation of Cascade Multilevel Inverter-based STATCOM", IETE Journal of Research, Vol. 56, No. 2, 2010, pp. 119-128.

[11] E. Babaei, "A Cascade Multilevel Converter Topology with Reduced Number of Switches", IEEE Transactions on Power Electronics, Vol. 23, No. 6, 2008, pp. 2657-2664.

[12] D. A. B. Zambra, C. Rech, J. R. Pinheiro, "Comparison of Neutral-Point-Clamped, Symmetrical, and Hybrid Asymmetrical Multilevel Inverters", IEEE Transactions on Industrial Electronics, Vol. 57, No. 7, 2010, pp. 2297-2306.

[13] M. F. Kangarlu, E. Babaei, S. Laali, "Symmetric multilevel inverter with reduced components based on non-insulated DC voltage sources", IET Power Electronics, Vol. 5, No. 5, 2012, pp. 571-581.

[14] E. Babaei, S. H. Hosseini, "New cascaded multilevel inverter topology with minimum number of switches", Energy Conversion and Management, Vol. 50, No. 11, pp. 2761–2767, Nov. 2009.

[15] N. A. Rahim, M. F. M. Elias, W. P. Hew, "Transistor-Clamped H-Bridge Based Cascaded Multilevel Inverter with New Method of Capacitor Voltage Balancing", IEEE Transactions on Industrial Electronics, Vol. 60, No. 8, 2013, pp. 2943-2956.

[16] E. Babaei, S. Alilu, S. Laali, "A new general topology for cascaded multilevel inverters with reduced number of components based on developed H-bridge", IEEE Transactions on Industrial Electronics, Vol. 61, No. 8, 2014, pp. 3932-3939.

[17] J. N. Chiasson, L. M. Tolbert, K. J. McKenzie, Zhong Du, "Control of a Multilevel Converter Using Resultant Theory", IEEE Transactions on Control Systems Technology, Vol. 11, No. 3, 2003, pp. 345-353.

[18] B. Ozpineci, L. M. Tolbert, J. N. Chiasson, "Harmonic Optimization of Multilevel Converters Using Genetic Algorithms", IEEE Power Electronics Letters, Vol. 3, No. 3, 2005, pp. 92-95.

[19] J. N. Chiasson, L. M. Tolbert, K. J. McKenzie, Z. Du, "A new approach to solving the harmonic elimination equations for a multilevel converter", Proceedings of the 38th IAS Annual Meeting on Conference Record of the Industry Applications Conference, Salt Lake City, UT, USA, 12-16 October 2003, pp. 640-645.

[20] J. Sun, S. Beineke, H. Grotstollen, "Optimal PWM based on real-time solution of harmonic elimination equations", IEEE Transactions on Power Electronics, Vol. 11, No. 4, pp. 612-621, Jul 1996.

[21] J. Kumar, B. Das, P. Agarwal, "Selective Harmonic Elimination Technique for a Multilevel Inverter", Proceedings of the15th National Power System Conference, Bombay, India, 16-18 December 2008, pp. 608-613.

[22] H. Taghizadeh, M. T. Hagh, "Harmonic Elimination of Cascade Multilevel Inverters with Nonequal DC Sources Using Particle Swarm Optimization", IEEE Transactions on Industrial Electronics, Vol. 57, No. 11, 2010, pp. 3678-3684.

[23] J. Kumar, B. Das, P. Agarwal, "Optimized Switching Scheme of a Cascade Multi-level Inverter", Electric Power Components and Systems, Vol. 38, No. 4, 2022, p. 445.

Obstructive Sleep Apnea Detection based on ECG Signal using Statistical Features of Wavelet Subband

Original Scientific Paper

Achmad Rizal
Telkom University,
School of Electrical Engineering
Telekomunikasi no.1, Bandung, Indonesia
achmadrizal@telkomuniversity.ac.id

Hilman Fauzi
Telkom University,
School of Electrical Engineering
Telekomunikasi no.1, Bandung, Indonesia
hilmanfauzitsp@telkomuniversity.ac.id

Sugondo Hadiyoso
Telkom University,
School of Applied Science
Telekomunikasi no.1, Bandung, Indonesia
sugondo@telkomuniversity.ac.id

Rahmat Widadi
Institut Teknologi Telkom Purwokerto,
Faculty of Telecommunication and Electrical Engineering
D.I. Panjaitan no.128, Purwokerto, Indonesia
rahmat@ittelkom-pwt.ac.id

Abstract *– One of the respiratory disorders is obstructive sleep apnea (OSA). OSA occurs when a person sleeps. OSA causes breathing to stop momentarily due to obstruction in the airways. In this condition, people with OSA will be deprived of oxygen, sleep awake and short of breath. Diagnosis of OSA by a doctor can be done by confirming the patient's complaints during sleep, sleep patterns, and other symptoms that point to OSA. Another way of diagnosing OSA is a polysomnography (PSG) examination in the laboratory to analyze apnea and hypopnea. However, this examination tends to be high cost and time consuming. An alternative diagnostic tool is an electrocardiogram (ECG) examination referring to changes in the mechanism of ECG-derived respiration (EDR). So digital ECG signal analysis is a potential tool for OSA detection. Therefore, in this study, it is proposed to classify OSA based on ECG signals using wavelets and statistical parameters. Statistical parameters include mean, variance, skewness kurtosis entropy calculated on the signal decomposition results. The validation performance of the proposed method is carried out using a support vector machine, k-nearest neighbor (k-NN), and ensemble classifier. The proposed method produces the highest accuracy of 89.2% using a bagged tree where all features are used as predictors. From this study, it is hoped that ECG signal analysis can be used to complete clinical diagnosis in detecting OSA.*

Keywords*: Obstructive sleep apnea, electrocardiogram, Wavelet transform, statistical parameter*

1. INTRODUCTION

An electrocardiogram (ECG) signal is a biomedical signal generated by bioelectric phenomena in the heart. The ECG signal describes the process of contraction of the heart muscle to pump blood out of the heart [1]. In addition to providing information about heart health, ECG signals can also provide other health information. One information the ECG signal can provide is the occurrence of obstructive sleep apnea (OSA) [2]. OSA is a disorder that occurs during sleep, where a block in the upper respiratory tract occurs during sleep [3]. The occurrence of OSA can be seen from the mechanism of ECG-derived respiration (EDR), where a disturbance in breathing will cause a change in the pattern of the ECG signal [4]. Changes to this ECG signal can be a change in signal orientation, a change in signal shape, or a change in signal rhythm.

OSA can be detected or monitored using several methods. The standard method used is Polysomnography (PSG) [5]. PSG is less practical because it requires monitoring throughout the night and uses a lot of devices. PSG is also only used for patients with OSA with a certain severity [2]. The following method is OSA monitoring using a camera that detects the patient's movement during sleep [6]. The advantage of this method is that the system is contactless but requires heavy processing to be carried out on video recordings over a long period. OSA detection using a camera can only be done if the subject makes specific movements in OSA conditions. Another method is to use the sound

of snoring as input. This method uses voice processing techniques to detect sleep apnea [7]. Because the voice frequency is relatively high, the sampling frequency will also be quite high, affecting the computational complexity. Sound signal processing for OSA detection is strongly influenced by environmental noise, for example, in patients with sleep talking. Other investigators have used a combination of ECG, electromyogram (EMG), and electroencephalogram (EEG) signals [8], or ECG and photoplethysmogram (PPG) signals by exploiting changes in oxygen saturation caused by OSA [9]. Simultaneous use of multiple signals for OSA detection creates complexity in installation and has the potential to cause subject discomfort. However, the use of a single ECG as an input signal to detect OSA has become the choice of many researchers because it is more practical with more straightforward computations [2], [10]–[12].

Various methods have been developed to detect OSA using ECG signals. OSA detection using ECG rhythm changes is usually initiated by seeing the R wave in the ECG signal. The cessation of breathing results in a difference in the distance between the R waves so that by calculating several parameters such as the standard deviation of the RR-interval, NN50 measure (type1 and type 2), pNN50, RMSSD, etc. [13][14][15]. The determination of the HRV of the ECG signal is highly dependent on the detection of the QRS of the ECG signal. If any QRS fails to detect, then the HRV parameter will change, which will cause misclassification. The short length of the ECG signal segmentation also affects the accuracy because if the ECG signal segment is too long, the HRV changes due to OSA cannot be detected [13]. Changes in the rhythm of this signal will result in a difference in the shape or orientation of the ECG signal. For this reason, several researchers use the morphology of the ECG signal as a feature for classifying normal and OSA ECG signals. The morphology of the ECG signal was measured using the entropy [16], the intrinsic mode function (IMF) of the ECG signal [17][18], or signal fluctuations in the wavelet subband [19]. The use of EMD in the decomposition of the ECG signal for OSA detection has the potential to eliminate the information contained in the ECG signal. As we know, in EMD, the local oscillation of the signal is removed to obtain the intrinsic mode function (IMF) of the signal [17].

The wavelet method has its advantages in processing ECG signals for OSA detection. The advantage is due to the flexibility of wavelets in decomposing the signal to obtain a subband with the desired information content. In this study, it is proposed to classify OSA on ECG signals using DWT and statistical parameters. The decomposition level equals 3, so it only produces four subbands. Meanwhile, for characteristics, statistical characteristics of order 1. are used. These statistical characteristics will display signal characteristics such as mean, variance, skewness, kurtosis, and entropy. The level of decomposition used in this study is less than that proposed in [19]. In [19], decomposition level 8 is used to obtain nine subbands. Meanwhile, the features used include several types of entropy (ApEn, FE, Shannon entropy, Correct Conditional Entropy), Poincare Plot Features, recurrent plot, mean absolute deviation, and variance. Although the proposed method is more straightforward than previous studies, It is hoped that this feature will be able to produce high accuracy in the classification of normal ECG and OSA.

2. MATERIAL AND METHODS

The method proposed in this study is shown in Figure 1. The recording of ECG signal is segmented into 1 minute of recording and then normalized. In the next stage, decomposition is carried out using the discrete wavelet transform (DWT) method to obtain the signal subband. The statistical parameters are calculated in each subband, which will be used as features. The last process is classification to determine whether OSA occurs or not. The details of each process will be explained in the following subsection.

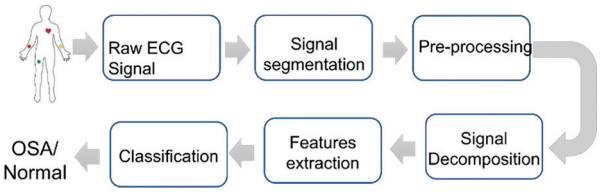


Fig. 1. The proposed method

2.1. DATA, SIGNAL SEGMENTATION, DAN PREPROCESSING

ECG signal data on OSA was obtained from Physionet, which consisted of 70 records, with 35 records being training data and 35 recordings of testing data [20][21]. Each record lasts 7-10 hours with a sampling frequency of 100 Hz. Each 1-minute segment or 6000 samples has been annotated by experts in the form of OSA or normal. In this study, the ECG signal data with and without OSA were combined for different subjects, so the method used was subject-independent.

In the segmentation signal, the normalization process is carried out using Eq. (1).

$$y = \frac{x - \bar{x}}{\sigma_x} \tag{1}$$

where y = ouput signal, x = input signal, $\bar{x}$ = mean of input signal, and σ_x = standard deviation of x.

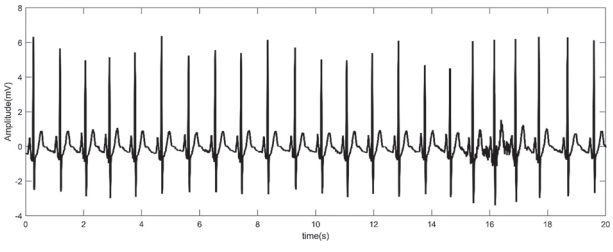


Fig. 2. The ECG signal in normal condition

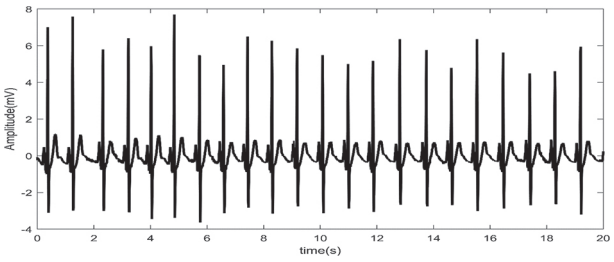


Fig. 3. The ECG signal with OSA condition

The number of data tested is 17010, with 10496 normal data and 6514 ECG data on OSA. Figure 2 and Figure 3 show examples of normal and OSA ECG signals.

2.2. WAVELET DECOMPOSITION

Wavelet transform aims to decompose the signal into several subbands, which are assumed to be stationary. The wavelet decomposition process involves a filtering process using a low-pass filter and an orthogonal high-pass filter and downsampling, which will reduce the number of signal samples by half [22]. This process will divide the signal into two subbands with energy that is maintained. In this study, Db2 is used as the mother wavelet with the decomposition level shown in Figure 4.

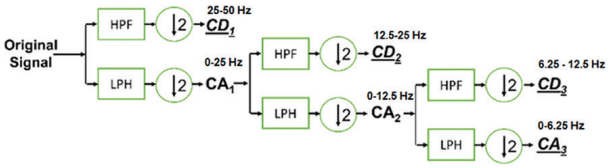


Fig. 4. Wavelet decomposition scheme which is implemented in this study

The bandwidth for each subband will be obtained in the ECG signal with a sampling frequency of 100 Hz, as shown in Table 1. In this study, the original signal is decomposed at three levels to obtain high-frequency information which is represented as detail coefficients (CD) and low-frequency information which is represented as approximation coefficients (CA). Most of the signal components will occupy the CA3 subband according to the characteristics of the ECG signal.

Table 1. Bandwidth of each subband

Subband	Band of frequency
CD1	25-50
CD2	12.5 – 25
CD3	6.25 – 12.5
CA3	0 – 6.25

2.3. FEATURE EXTRACTION

The result of wavelet decomposition is calculated statistical parameters as a signal feature in each subband. The statistical characteristics include mean, variance, skewness, kurtosis, and entropy. The five statistical features can be calculated using a mathematical formula as in Eq. 2 – 6.

$$Mean\ (\bar{Y}) = \frac{\sum_{i=1}^{N} Y_i}{N} \quad (2)$$

$$VAR = \frac{1}{N-1}\sum_{i=1}^{N} {Y_i}^2 \quad (3)$$

$$Skewness = \frac{\sum_{i=1}^{N}(Y_i-\bar{Y})^3/N}{s^3} \quad (4)$$

$$Kurtosis = \frac{\sum_{i=1}^{N}(Y_i-\bar{Y})^4/N}{s^4} \quad (5)$$

$$H = -\sum_{i=1}^{N} P_i \log P_i \quad (6)$$

The five statistical parameters are expected to differentiate between normal ECG signals and ECG signals under OSA conditions.

2.4. CLASSIFICATION

A. Support Vector Machine (SVM)

SVM is a learning system used to test linear function hypotheses in high-dimensional feature space; the computer will be trained using an algorithm based on optimization theory and statistical learning theory [23]. SVM can work on non-linear data by applying a kernel approach to the data set's initial features. Lower dimensions are mapped to higher dimensions by kernel functions [24].

B. K-nearest neighbor (KNN)

K-nearest neighbor (KNN) is a distance-based classification method that determines classification results based on the largest training data class with the closest distance to test data [25]. The following are the steps of the K-Nearest Neighbor algorithm:

1. Determine the value of the parameter K; K is the number of nearest neighbors.
2. Determine the distance between the test data and all existing data in the training data.
3. Sort the distances and find the nearest neighbor based on the shortest distance to the K data.
4. Determine the closest neighbor's category.

The Euclidean distance was used in this study as in Eq. (7). In this study, various variations of KNN were used to obtain the highest accuracy.

$$D_e(xi, yi) = \sqrt{(x1-y1)^2 + (x2-y2)^2 + \cdots + (xp-yp)^2} \quad (7)$$

C. Ensemble Classifier

The ensemble classifier is a data-driven approach to improving class balance. The goal of the ensemble algorithm approach is the same: to improve the algorithm without changing the data. There are two approaches that can be taken: the data level approach and the algorithm level approach [26]. Boosting and bagging are two

popular ensemble algorithms. AdaBoost is a classification algorithm that improves classification performance. At the same time, bagging is a simple but effective ensemble method that has been used in many real-world applications to improve the accuracy of classification algorithms. The results outperform random sampling [27].

D. N-Fold Cross-Validation

In this study, the N-fold cross-validation (NFCV) was used to avoid overfitting in the classification accuracy testing process. In NFCV, the data is divided into N datasets, with 1 data set used as the test while the N-1 dataset is used as the training data. The testing process is carried out N times so that each data set has been used as test data. Accuracy is taken from the average of each test. Because the classifiers used in this study are all supervised learning, NFCV is appropriate in this case. This study used N = 10.

3. RESULTS AND DISCUSSION

Figure 5 and Figure 6 show the decomposition results of normal ECG and ECG signals with OSA. From the number of signal samples, it can be seen that on CD1, the number of signal samples is half of the original signal, meanwhile, on CD2, the number of signal samples is half that of CD1, and the number of signal samples on CD3 and CA3 each has half the number of signal samples of CD2. This result occurs because there is a downsampling of 2 processes at each decomposition level, as shown in Figure 4. There is no difference between the decomposition results of normal ECG signals and ECG signals with OSA. For this reason, the feature extraction process is carried out by calculating statistical features in each subband, such as mean, variance, kurtosis, skewness, and entropy.

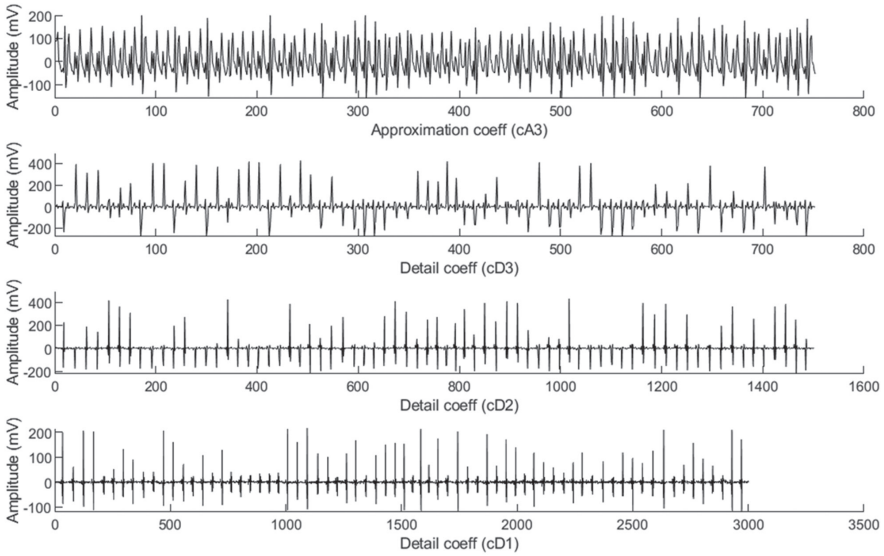


Fig. 5. Decomposition result of 60 s normal ECG from Fig. 2

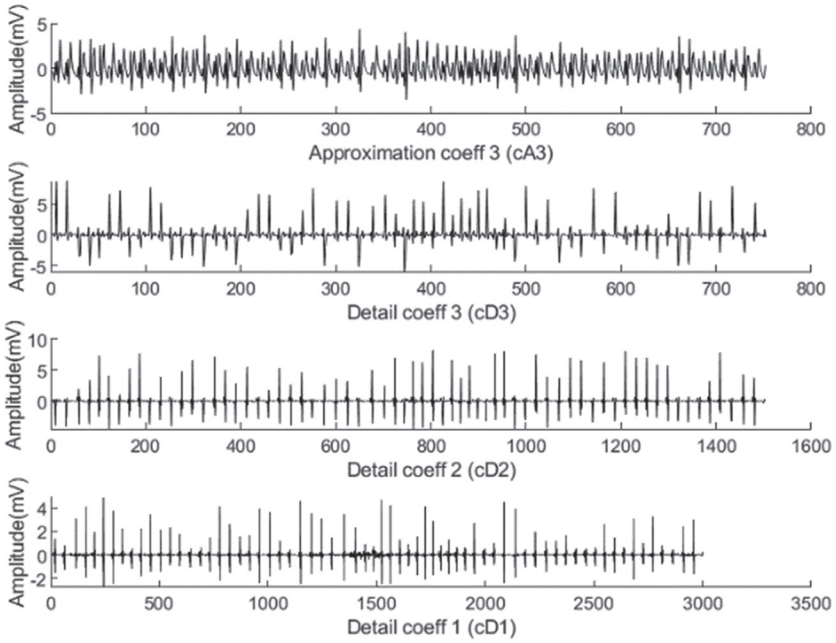


Fig. 6. Decomposition result of 60 s ECG with OSA from Fig 3

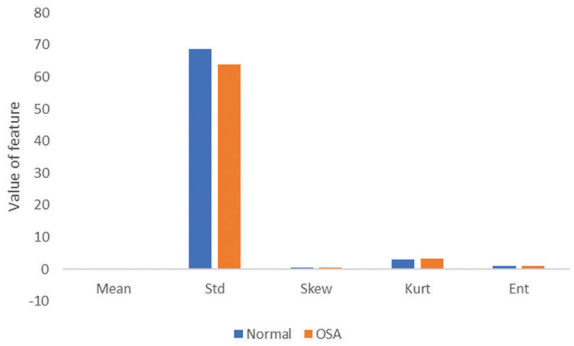


Fig. 7. Generated features from CA3

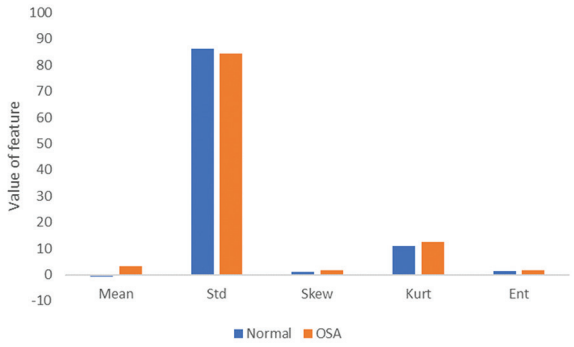


Fig. 8. Generated features from CD3

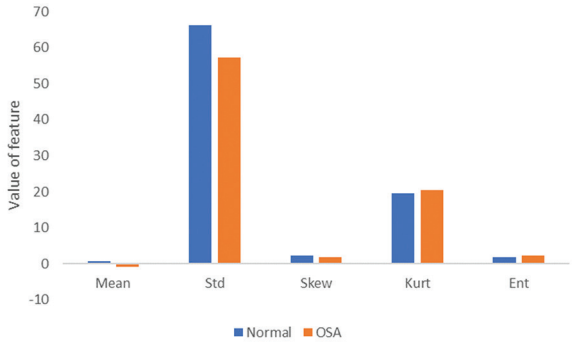


Fig. 9. Generated features from CD2

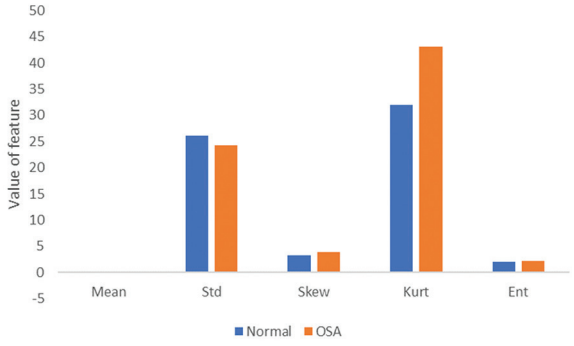


Fig. 10. Generated features from CD1

Figure 7 - Figure 10 show the statistical characteristics generated for each subband. Some of the resulting characteristic values are small enough that they are insignificant compared to other features. The mean feature looks very small because the mean value of the signal is close to zero due to the normalization process. Meanwhile, the standard deviation value tends to produce a reasonably significant value, and there is a tendency for the standard deviation of the normal ECG signal to be greater than the standard deviation of the OSA ECG. Kurtosis on a normal ECG tends to be lower than on an OSA ECG. Based on the distribution of the resulting characteristics, the accuracy will tend to be high even though it will not reach 100% because some values tend to be the same as entropy.

Table 2. Accuracy using SVM

	all feature	CA3	CD3	CD2	CD1
Linear SVM	75.1	61.7	62.7	67.4	61.7
Quadratic SVM	85.3	55.7	59.2	69.5	50.2
Cubic SVM	82.1	46.6	39.9	53.2	39
Fine Gaussian SVM	86.5	79.1	84.5	84.4	81.5
Medium Gaussian SVM	84.6	74.6	80.9	77.4	74.6
Coarse Gaussian SVM	78.4	61.8	76.2	72	64.4

Table 3. Accuracy using K-NN

	all feature	CA3	CD3	CD2	CD1
Fine KNN	82.9	77.7	79.4	79.7	79.7
Medium KNN	84.8	79.8	83.4	83	82.3
Coarse KNN	82.9	76.9	80.9	79.8	78.5
Cosine KNN	83.9	77.4	80.5	80.2	78.8
Cubic KNN	83.6	79.3	83.3	82.6	82
Weighted KNN	85.4	80.3	83.4	83	82.6

Table 4. Accuracy (%) using Ensemble Classifiers

	all feature	CA3	CD3	CD2	CD1
Boosted trees	83.1	78.8	80.5	78.4	76.6
Bagged trees	89.2	81.5	83.6	84.6	84.4
Subspace discriminant	64.3	61.5	60.6	62.3	62.2
Subspace KNN	87.1	76.9	78.3	77.6	78
RUSBoosted Trees	82.5	70.7	78.7	76.9	72.4

Table 2 - Table 4 show the classification accuracy using SVM, KNN, and ensemble classifier. Table 2 presents that the highest accuracy is 86.5% using all features with fine gaussian SVM as a classifier. This accuracy value is better than only using features in one subband, 84.5%, with the same SVM features on the CD3 subband.

Table 3 shows the highest accuracy of 85.4% generated by the overall characteristics with weighted KNN as a classifier. One subband yields the highest accuracy of 83.4% using CD3. Meanwhile, the ensemble classifier in Table 4 produces the highest accuracy of 89.2% using bagged trees and overall characteristics. Meanwhile, one subband only has the highest accuracy of 84.6% on CD2.

Individual CD3 subbands tend to produce the highest accuracy compared to other subbands because significant changes due to OSA occur in the CD3 subband, which occupies a frequency of 6.25 – 12.5 Hz. This is because the change in heart rate due to OSA is not very significant.

A significant change is precisely in the regularity of the heart rate and the sudden change in orientation that occurs in recording the ECG signal.

The difference between a normal ECG and an OSA ECG is in the sudden change in heart rate or orientation due to apnea. This study's 1-minute recording time will display 60-80 QRS for each data plot. The pattern produced by the OSA condition mainly affects only 1-2 QRS patterns. This condition generally does not provide a significant change in the pattern in the frequency domain. Several studies cut the ECG recording to 30 seconds with the consequence of adding a label to the data cut-off [28].

Table 5. Comparison with previous research

Author	Method	Accuracy (%)
Rizal, Iman, & Fauzi [16]	Multiscale entropy, SVM	85.6
Rizal, Siregar, & Fauzi [13]	Heartrate variability, SVM	89.5
Zhu et al. [29]	Plot of pointcare & recurrent, entropy, SVM	94.63
Razi, Einalou, & Manthouri [15]	LDA, random forest	95.1
Proposed method	Wavelet transform, first order statistical, bagged trees	89.2

Table 5 presents a comparison between the proposed method and previous studies. Multiscale entropy produces the highest accuracy of 85.6% [16]. HRV used in other studies yielded the highest accuracy of 89.5%. Meanwhile, LDA and random forest as a classifier produce the highest accuracy of 95.1% [15]. Even though it does not produce higher accuracy than previous research, the proposed method has several advantages, including a simple process, and still leaves some development possibilities. If, in this study, the statistical parameters were calculated in the subband, then in the following research, the use of entropy, fractal dimensions, or other parameters can be explored [30]. In addition, in this study, only level 3 decomposition was used; the exploration of selecting the right subband as a feature of the OSA ECG becomes an exciting topic in future research. The proposed method is expected to be an alternative feature extraction method in the ECG OSA classification.

4. CONCLUSION

This study proposes a method for OSA detection based on ECG signal. In the feature extraction stage, the ECG signal is decomposed into three subbands using a wavelet to calculate the statistical parameters for each subband. Statistical parameters calculated include mean, variance, skewness, kurtosis, and entropy. The proposed method was then tested on the OSA ECG dataset taken from Physionet. In the classification stage, the number of data tested is 17010 consisting of 10496 normal and 6514 OSA. Several feature usage scenarios and classifiers are applied to find the best classification performance. The classifiers used in this study are SVM, KNN and ensembles with various variations of the trick kernel. From all test scenarios, the highest accuracy is obtained if all features are used as predictors. The highest accuracy achieved was 89.2% using bagged trees. This study is expected to be a reference method in the detection of OSA ECG so that it can complete the clinical diagnosis. This study shows that ECG signal analysis can be an alternative tool in OSA detection considering the low cost, simple to manage, and is closely related with OSA compared to the PSG examination. Future studies still offer opportunities to explore feature extraction methods especially time domain based methods and other classifiers to improve detection accuracy. In addition, the detection method using a deep learning approach is also interesting to study.

5. REFERENCES:

[1] Z. Said, Y. el Hassouani, "A New Approach for Extracting and Characterizing Fetal Electrocardiogram", Traitement du Signal, Vol. 37, No. 3, 2020, pp. 379-386.

[2] O. Faust, U. R. Acharya, E. Y. K. Ng, H. Fujita, "A review of ECG-Based diagnosis support systems for obstructive sleep apnea", Journal of Mechanics in Medicine and Biology, Vol. 16, No. 1, 2016.

[3] J. Zhu, A. Zhou, Q. Gong, Y. Zhou, J. Huang, Z. Chen, "Detection of Sleep Apnea from Electrocardiogram and Pulse Oximetry Signals Using Random Forest", Applied Sciences, Vol. 12, No. 9, 2022, p. 4218.

[4] A. A. Khatami, M. A. C. Sebayang, A. Rizal, D. T. Barus, "Obstructive Sleep Apnea Detection using ECG Signal: A Survey", Technology Reports of Kansai University, Vol. 62, No. 4, 2020, pp. 1267-1274.

[5] B. Jafari, V. Mohsenin, "Polysomnography", Encyclopedia of Sleep, Academic Press, 2013.

[6] M. van Gastel, S. Stuijk, S. Overeem, J. P. van Dijk, M. M. van Gilst, G. de Haan, "Camera-Based Vital Signs Monitoring During Sleep - A Proof of Concept Study", IEEE Journal of Biomedical and Health Informatics, Vol. 25, No. 5, 2021, pp. 1409-1418.

[7] J. Solé-Casals et al., "Detection of severe obstructive sleep apnea through voice analysis", Applied Soft Computing Journal, Vol. 23, 2014, pp. 346-354.

[8] M. K. Moridani, M. Heydar, S. S. Jabbari Behnam, "A Reliable Algorithm Based on Combination of EMG,

ECG and EEG Signals for Sleep Apnea Detection: (A Reliable Algorithm for Sleep Apnea Detection)", Proceedings of the 5th Conference on Knowledge Based Engineering and Innovation, Tehran, Iran, Feb. 2019, pp. 256-262.

[9] M. K. Uçar, M. R. Bozkurt, C. Bilgin, K. Polat, "Automatic detection of respiratory arrests in OSA patients using PPG and machine learning techniques", Neural Computing and Applications, Vol. 28, No. 10, 2017, pp. 2931-2945

[10] G. B. Papini et al. "On the generalizability of ECG-based obstructive sleep apnea monitoring: merits and limitations of the Apnea-ECG database", Proceedings of the 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Honolulu, Hawaii, USA, July 2018, pp. 6022-6025.

[11] V. P. Rachim, G. Li, W. Y. Chung, "Sleep apnea classification using ECG-signal wavelet-PCA features", Bio-Medical Materials and Engineering, Vol. 24, No. 6, 2014, pp. 2875-2882.

[12] V. Ankitha, P. Manimegalai, P. S. H. Jose, P. Raji, "Literature Review on Sleep Apnea Analysis by Machine Learning Algorithms Using ECG Signals", Journal of Physics: Conference Series, Vol. 1937, No. 1, 2021, p. 012054.

[13] A. Rizal, F. D. A. A. Siregar, H. T. Fauzi, "Obstructive Sleep Apnea (OSA) Classification Based on Heart Rate Variability (HRV) on Electrocardiogram (ECG) Signal Using Support Vector Machine (SVM)", Traitement du Signal, Vol. 39, No. 2, 2022, pp. 469-474.

[14] D. S. Sisodia, K. Sachdeva, A. Anuragi, "Sleep order detection model using support vector machines and features extracted from brain ECG signals", Proceedings of the International Conference on Inventive Computing and Informatics, Coimbatore, India, 2018, pp. 1011-1015.

[15] A. P. Razi, Z. Einalou, M. Manthouri, "Sleep Apnea Classification Using Random Forest via ECG", Sleep Vigilance, Vol. 5, 2021, pp. 141-146.

[16] A. Rizal, U. R. Iman, H. Fauzi, "Classification of Sleep Apnea using Multi Scale Entropy on Electrocardiogram Signal", Journal of Online and Biomedical Engineering, Vol. 17, No. 14, 2021, pp. 79-89.

[17] M. O. Mendez et al., "Automatic screening of obstructive sleep apnea from the ECG based on empirical mode decomposition and wavelet analysis", Physiological Measurement, Vol. 31, No. 3, 2010, pp. 273-289.

[18] D. S. B. Sundaram, R. Balasubramani, S. Shivaram, A. Muthyala, S. P. Arunachalam, "Single Lead ECG Discrimination between Normal Sinus Rhythm and Sleep Apnea with Intrinsic Mode Function Complexity Index Using Empirical Mode Decomposition", Proceedings of the IEEE International Conference on Electro Information Technology, Rochester, MI, USA, 2018, pp. 719-722.

[19] A. Zarei, B. M. Asl, "Automatic Detection of Obstructive Sleep Apnea Using Wavelet Transform and Entropy-Based Features From Single-Lead ECG Signal", IEEE Journal of Biomedical and Health Informatics, Vol. 23, No. 3, 2019, pp. 1011-1021.

[20] A. L. Goldberger et al., "PhysioBank, PhysioToolkit, and PhysioNet Components of a New Research Resource for Complex Physiologic Signals", Ciculation, Vol. 101, 2000, pp. e215-e220.

[21] T. Penzel, G. B. Moody, R. G. Mark, A. L. Goldberger, J. H. Peter, "The apnea-ECG database", Computers in Cardiology 2000, Vol. 27, No. 1, 2000, pp. 255-258.

[22] A. Sharmila, V. N. Sharma, G. Dey, K. Yashwanth, "Detection of Artial Fibrillation Disorder by ECG Using Discrete Wavelet Transforms", World Applied Sciences Journal, Vol. 33, No. 5, 2015, pp. 752-756.

[23] C. Cortes, V. Vapnik, "Support vector machine", Machine Learning, Vol. 20, 1995, pp. 1303-1308.

[24] B. Abbaszadeh, T. Haddad, M. C. E. Yagoub, "Probabilistic prediction of Epileptic Seizures using SVM", Proceedings of the 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Berlin, Germany, July 2019, pp. 3442-3445.

[25] Y. N. Fuadah, A. W. Setiawan, T. L. R. Mengko, "Performing High Accuracy of The System for Cataract Detection Using Statistical Texture Analysis and K-Nearest Neighbor", Proceedings of the International Seminar on Intelligent Technology and Its Applications Performing, Surabaya, Indonesia, 2015, pp. 85-88.

[26] V. Murugananthan, U. K. Durairaj, "RUS boost tree ensemble classifiers for occupancy detection", International Journal of Recent Technology and Engineering, Vol. 8, No. 2 Special Issue 2, 2019, pp. 272-277.

[27] X. Yan, M. Jia, "Intelligent fault diagnosis of rotating machinery using improved multiscale dispersion entropy and mRMR feature selection", Knowledge-Based Systems, Vol. 163, 2019, pp. 450-471.

[28] L. Almazaydeh, K. Elleithy, M. Faezipour, "Obstructive sleep apnea detection using SVM-based classification of ECG signal features", Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS, San Diego, CA, USA, 2012, pp. 4938-4941.

[29] H. Zhu, C. H. Koo, C. K. Wu, W. H. Wan, Y. T. Tsang, K. F. Tsang, "Sleep apnea monitoring for smart healthcare", Proceedings of the 44th Annual Conference of the IEEE Industrial Electronics Society, Washington, DC, USA, 2018, pp. 4726-4729.

[30] A. Rizal, A. Puspitasari, "Lung Sound Classification Using Wavelet Transform and Entropy to Detect Lung Abnormality", Serbian Journal of Electrical Engineering, Vol. 19, No. 1, 2022, pp. 79-98.

Color Image Encryption Using LFSR, DNA, and 3D Chaotic Maps

Original Scientific Paper

Salah Taha Allawi
Mustansiriyah University,
Department of Computer Science, College of Science
Baghdad, Iraq
salah.taha@uomustansiriyah.edu.iq

Dina Riadh Alshibani
Mustansiriyah University,
Department of Computer Science, College of Science
Baghdad, Iraq
dinashibani@uomustansiriyah.edu.iq

Abstract *– One of the most important challenges facing researchers is to find new methods to protect data sent over the Internet and prevent unauthorized access to it. In this paper, we present a new method for encrypting image data divided into two stages. The first stage requires redistributing the positions of the pixels by using a key of random numbers generated by linear feedback shift registers and then encrypting the data using deoxyribonucleic acid rules. The data generated in the previous stage is encrypted again using chaotic maps to increase the level of security in the second stage. Several statistical tests were implemented to verify the efficiency of the proposed method and compare the results with the work of other researchers. The results of the tests proved a reasonable safety rate compared to other techniques.*

Keywords*: DNA encoding, LFSR, 3D chaotic maps, Encrypted image*

1. INTRODUCTION

Currently, there are many ways to transfer and store information. While modern technology affords easy transmission of information, many risks affect the security and safety of this information [1]. Traditional encryption systems, such as DES and AES, are effective and secure encryption systems when dealing with text data. However, they are not effective when dealing with images because of the characteristics of images, such as redundancy and strong correlation between adjacent pixels [2], [3].

The efficiency, strength, and complexity of the algorithm used to encrypt data are based on the data type. Since images are unique type of data, the algorithm used to encrypt them must be complex [4]. These reasons prompted researchers to discover new methods to protect image data that is transmitted through unsecured networks. One of the essential methods in this field is using chaotic maps to encrypt image data [5]. The scientist Lorenzo is the founder of chaos theory, which has various advantages when applied to image encryption. Chaotic maps are suitable for designing image-encryption systems. During the last decade, various studies and research in this field have been published [6]. As a result of scientific advancement in the field of image encryption, DNA computing was introduced [7]. Recent studies have proved that deoxyribonucleic acid (DNA) coding technology can effectively resist chosen plain-text attacks, improving the security of cryptosystems [8], [9]. Many image-encryption schemes use DNA technology because of its advantages: ultra-low energy consumption, large storage density, and high parallelism [10]. New methods of encrypting information using DNA provide increased security through the use of the biological structure of DNA, which is a natural carrier of information in binary form, by encoding information according to the four bases of DNA (T-11, C-01, G-10, and A-00) [11], [12]. Researchers have recently merged chaotic mapping and DNA technology to create more effective image-encryption methods [13].

In this paper, we provide a literature review on a selection of previous research on image encryption using DNA (Section 2). Then, chaotic maps and DNA encoding are explained in Sections 3 and 4, respectively. Sections 5 and 6 outline the proposed method and experiments to test the image-encryption system. Section 7 explains the test results and makes comparisons with the results of other methods. Finally, we discuss the conclusions in Section 8.

2. LITERATURE REVIEW

This section reviews some methods that use DNA to encode image data.

In [4], the authors proposed a new method comprising of two stages for encrypting color images based on 3D chaotic maps and DNA coding. In the first stage, the initial step is to generate a key using a 3D Arnold map and encode the image data and the key according to the DNA coding rule. The second step involves using the same key to reopen the image data codes and redistribute the image pixel positions after applying the XOR between the key and image codes. In the second stage, a 3D logistic map encrypts image data by generating three data-encryption keys.

In [14], the authors proposed a new method to encrypt images that combined chaotic maps and DNA to produce an algorithm that provides security and protection for confidential data when sent.

In [15], the authors proposed a new method to encrypt and confuse color images based on the DNA sequence with a new Beta chaotic map. First, the DNA addition operation is used to diffuse each component of the original image, and then a new Beta chaotic map shuffles the resulting image. Finally, to produce the encrypted image, a DNA XOR operation is applied between the shuffled DNA image and the key generated by the two new proposed chaotic maps, Beta and Sine chaotic maps.

In [16], the authors proposed a new method for encoding color images that combined a hybrid chaotic map and DNA. This method first requires dividing each image color into n parts and then rearranging the parts of the image. A Chaotic hybrid map redistributes the pixel positions of these parts, and then the parts are grouped to form the encrypted image. Finally, the image data is encoded using DNA technology.

In [17], the authors proposed a new method of encrypting images based on DNA sequence encryption and an enhanced 2D logistic Sine chaotic map (2D-LSMM). The logistic map is used to control the input of the Sine map, and the 2D-LSMM chaotic sequences are used to determine the operation and encoding rules of DNA sequences.

In [18], the authors proposed a new method that presented a hybrid model that uses the Lorenz–Rossler chaotic map to encode color images. The random sequences that are generated using Lorenz–Rossler chaotic systems are used to encode the essential components (red, green, and blue [RGB] channels) of the color image. To encode the original image, they used the rules of a DNA encoding system.

3. CHAOTIC SYSTEMS

Chaos theory is a branch of mathematics that is based on nonlinear and deterministic behavior. Any change in control parameters or initial values leads to a change in the chaotic output, which means chaotic systems have a higher sensitivity to their initial conditions. Chaotic systems are used in various areas, such as data encryption, data hiding, watermarking, and other areas that require unexpected results and outputs [19], [20]. Simplicity is a characteristic of a logistic map. Equation 1 explains the chaotic behavior of a 1D logistic map [21].

$$P_{i+1}=\mu * P_i * (1-P_i) \quad (1)$$

Where (μ) is the system control parameter, (P_0) is the initial state, (i) is the number of iterations, and the value of (P_{i+1}) is between (0 and 1) for all (i). The value of the control parameter (μ) is between (0 and 4), and the best result occurs when it is closer to 4. A 3D logistic map is better than a 1D logistic map [22]. Equations 2, 3, and 4 explain the chaotic behavior of a 3D logistic map.

$$x_{i+1}=\alpha x_i (1-x_i)+\beta y_i^2 x_i+\gamma z_i^3 \quad (2)$$

$$y_{i+1}=\alpha y_i (1-y_i)+\beta z_i^2 y_i+\gamma x_i^3 \quad (3)$$

$$z_{i+1}=\alpha z_i (1-z_i)+\beta x_i^2 z_i+\gamma y_i^3 \quad (4)$$

Where (β, α, γ) are three parameters ($0 < \beta < 0.022$), ($0 < \gamma < 0.015$), ($3.53 < \alpha < 3.81$), and z_0, y_0, x_0 take a value between (0 and 1).

4. DNA ENCODING

Information in DNA is stored as a code comprising four chemical bases: guanine (G), adenine (A), thymine (T), and cytosine (C). The rules are that "A" corresponds to "00", "C" corresponds to "01", "G" corresponds to "10", and "T" corresponds to "11". As in binary math, 1 and 0 are complements, so 11 and 00 are also complements. Similarly, 10 and 01 are complements. There are 24 types of coding groups, however, we use only 8 of them. Table 1 summarizes these rules [23]–[26].

Table 1. The rules of DNA encoding

	Rule 1	Rule 2	Rule 3	Rule 4	Rule 5	Rule 6	Rule 7	Rule 8
A	00	00	11	11	10	01	10	01
T	11	11	00	00	01	10	01	10
C	10	01	10	01	00	00	11	11
G	01	10	01	10	11	11	00	00

Many researchers used algebraic operations like subtraction, addition, and XOR due to the rapidly developing nature of DNA computations. Tables 2–5 explain the results of applying XOR with the rules (1, 3, 5, 7).

Table 2. Result applying XOR operation on DNA rule 1

Pixel	Key			
	A 00	T 11	C 10	G 01
A 00	A	T	C	G
T 11	T	A	G	C
C 10	C	G	A	T
G 01	G	C	T	A

Table 3. Result applying XOR operation on DNA rule 3

Pixel	Key			
	A 11	T 00	C 10	G 01
A 11	T	A	G	C
T 00	A	T	C	G
C 10	G	C	T	A
G 01	C	G	A	T

Table 4. Result applying XOR operation on DNA rule 5

Pixel	Key			
	A 10	T 01	C 00	G 11
A 10	C	G	A	T
T 01	G	C	T	A
C 00	A	T	C	G
G 11	T	A	G	C

Table 5. Result applying XOR operation on DNA rule 7

Pixel	Key			
	A 10	T 01	C 11	G 00
A 10	G	C	T	A
T 01	C	G	A	T
C 11	T	A	G	C
G 00	A	T	C	G

5. PROPOSED METHOD

The method proposed in this paper has two stages: encryption and decryption. Fig. 1 explains the general layout of the proposed method.

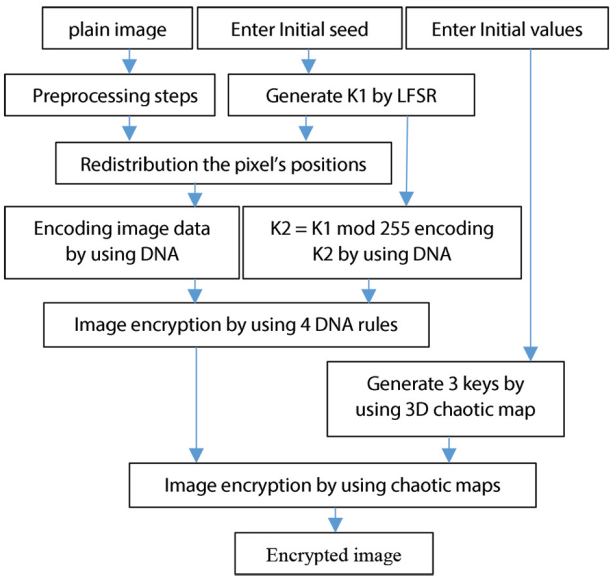


Fig. 1. The general layout for the proposed method

5.1. ENCRYPTION STAGE

The encryption stage aims to remove the correlation between adjacent pixels first by redistributing their positions and then changing the data values of the pixels by encrypting them using the rules of DNA coding and 3D chaotic maps.

5.1.1. PIXEL POSITION REDISTRIBUTED

In this stage, the original image's colors are separated into their essential RGB components, and then each color is divided into two equal parts (P1, P2). Using LFSR, generating a random key (K1) with non-repeated numbers of size (SK = H * W/2, where H and W represent the height and width of the image) and use the key to redistribute the positions of the pixels for each part (P1, P2). The random key is generated by using three registers with lengths of 29, 31, and 37 and join functions (2, 7, 13), (1, 2, 3), (1, 2, 3, 4, 5), respectively, and different initial values for each register. Fig. 2 explains the lengths and join functions for the registers.

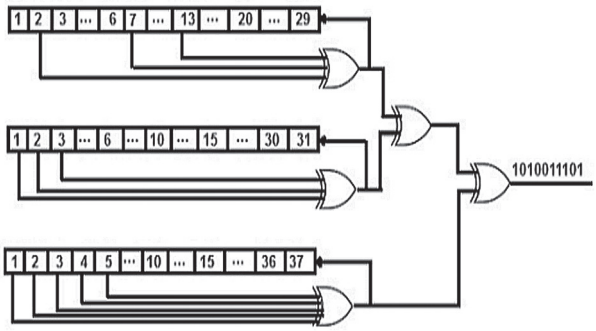


Fig. 2. Simplified drawing of a LFSR.

5.1.2. IMAGE ENCRYPTION USING DNA AND LFSR

The data for each color (P1, P2) is first converted to a binary value and then encoded using the first rule of DNA (A = 00, T = 11, C = 10, G = 01).

For example, if a pixel's value is (228), the binary number will be (11100100), rule 1 will be applied, and the outcome will be (TGCA).

A new key is calculated in the second step (K2 = K1 mod 256), then encoded using the first rule of DNA. The third step involves applying an XOR operation between the key (K2) values and the image data (P1, P2) that are coded to produce the encrypted image.

For example, when applying the XOR between the first code from the key (T) and the pixel (A) using Table 2, the result is (T). When applying the XOR between the second code from the key (C) and the pixel (G) using Table 3, the result is (A). This work continues until all image data is encrypted.

Algorithm 1 explains image data encryption using the rules of DNA and the LFSR.

Algorithm 1: The process of image encryption using the rules of DNA encoding and LFSR

Input: Plain color image
Output: Image encrypted by DNA encoding and LFSR

1. Enter an original color image (H * W).
2. Isolate image colors to essential RGB components.
3. Divide each color into two equal parts (P1, P2).
4. Convert data for each part (P1, P2) into a 1D array.
5. Generating a key (K1) with random non-duplicate numbers using LFSR with size (SK = H * W/2).
6. Redistribute pixel positions for each part (P1, P2) using the random key (K1).
7. Encode the data for each part (P1, P2) using the rule of DNA coding (rule 1) after converting it to a binary number.
8. Calculate a new key (K2 = K1 mod 256) with values ranging from (0 to 255).
9. Encode the key (K2) using the rule for coding DNA (rule 1) after converting it to a binary number.
10. Apply an operation XOR between the key and the data for each part (P1, P2) by using four DNA encoding rules (1, 3, 5, 7).

5.1.3. IMAGE ENCRYPTION USING 3D CHAOTIC MAPS

To encode the image data, generate three keys (Kr, Kg, Kb) by using 3D chaotic maps (refer to Equations 2, 3, 4), where one key is used to encrypt data relating to a specific color.

For example, encrypt the data for the color red (P1, P2) using the key (Kr).

Finally, the colors are combined to get the encoded image after the locations of the two parts (P1, P2) change for each color.

Algorithm 2 explains the process of image encryption using 3D chaotic maps.

Algorithm 2: The process of image encryption using 3D chaotic maps

Input: Image encrypted by DNA and LFSR
Output: Image encrypted
(result of the proposed method)

1. Generate 3 keys (Kr, Kg, and Kb) using the 3D chaotic map equations.
2. Apply the XOR operation between the keys and the color values.
3. Convert each part (P1, P2) into a 2D array.
4. Change the location of the two parts (P1, P2).
5. Reconstruct the image colors.
6. The result is an encrypted image.

5.2. DECRYPTION STAGE

Recovering the original image includes two stages. The first stage comprises entering the encrypted image, isolating the primary colors for the image (RGB), and dividing each color into two equal parts (P1, P2). Then, the data for each color is decrypted by generating the same three keys using 3D chaotic maps.

The second stage comprises the decryption of the data using the DNA rules (1,3,5,7) and the key generated by LFSR first. Second, recover the original position for each pixel using the key generated by LFSR. Algorithm 3 describes the decryption stage for images.

Algorithm 3: Process of image decryption

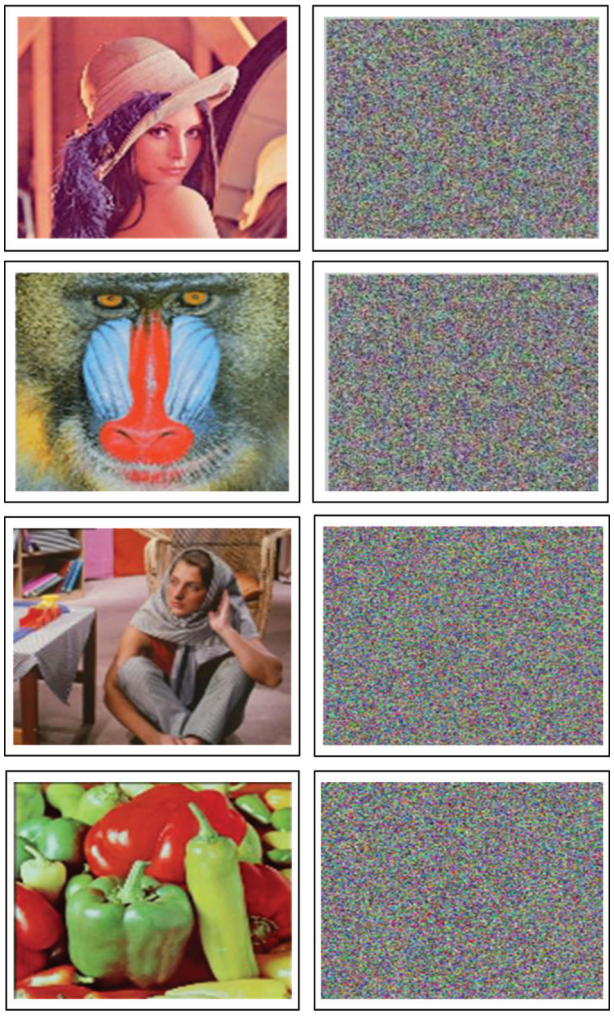
Input: Encrypted image
Output: Original color image

1. Enter the encrypted image (H * W).
2. Isolate the image colors into their essential RGB components.
3. Divide each color into two equal parts (P1, P2).
4. Convert the data for each part into a 1D array.
5. Generate 3 keys (Kr, Kg, and Kb) using 3D chaotic maps.
6. Apply the XOR operation between the keys and the color values.
7. Convert the data for each part (P1, P2) into a binary number and then encode it using the rule of DNA coding (rule 1).
8. Generate a key (K1) with random non-duplicate numbers using LFSR with size (SK = H * W/2).
9. Calculate a new key (K2 = K1 mod 256) with values ranging from (0 to 255).
10. Convert the key (K2) values to a binary number and then encode it using the rule for coding DNA (rule 1).
11. Apply an XOR operation between the key (K2) and each part (P1, P2) by using four DNA coding rules (1, 3, 5, 7).
12. Restore the original pixel positions using the key (K1).
13. Convert each part (P1, P2) into a 2D array.
14. Change the location of the parts (P1, P2).
15. Reconstruct the image colors.
16. The resulting image is the original image.

6. EXPERIMENTS

The proposed method for testing used a group of images (Lena, Peppers, Baboon, and Barbara) that were (256 * 256) in size. Fig. 3a shows the images used in the tests. The random number is generated using a group of bits generated through LFSR. The length of a series of bits depends on the number count required for the key.

The initial parameters $\beta = 0.02$, $\gamma = 0.015$, $\alpha = 3.84$, $y0 = 0.67$, $z0 = 0.97$, and $x0 = 0.97$ were used to produce the best results in the suggested method at the stage of data encoding using 3D chaotic maps. The proposed method's results are shown in Fig. 3 (b).



(a) (b)

Fig. 3. (a) Original images (b) Encrypted images

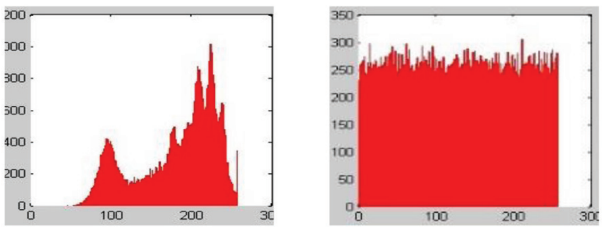
7. RESULTS AND DISCUSSION

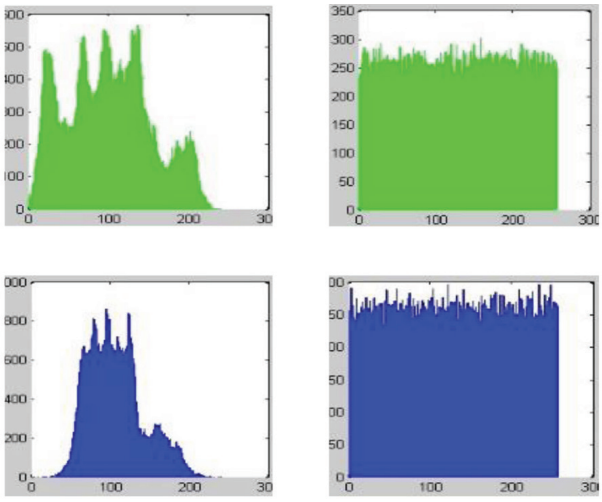
This section will present the results of applying various tests to the proposed method.

7.1. HISTOGRAM ANALYSIS

A histogram is employed to determine how pixels are arranged in an image and to learn more about the image's features. One method for attackers to discover an original image is to get the histogram of the encrypted image [27].

Fig. 4 (a) shows the original RGB histogram for the Lena image, and Fig. 4 (b) shows the RGB histogram for the encrypted Lena image.





(a) (b)

Fig. 4. RGB histogram analysis for the Lena image (a) Original image (b) Encrypted image

7.2. ENTROPY TEST

Entropy measures the strength and level of randomness provided by the encrypting system in terms of the amount of disorganization achieved in the image. Equation 5 is used to measure entropy [28].

$$H(q) = -\sum_{i=0}^{255} P(q_i)\ log_2\ P\ (q_i) \quad (5)$$

Where P(qi) is the probability of (qi) of pixels; the best value for the entropy of the encrypted image is the one closest to 8. Table 6 shows the entropy values of the images used in the experiments. The results show that the proposed method has excellent results and is near 8.

Table 6. Entropy values for the experiment's images.

image	Original image	Encrypted image
Lena	7.326	7.9973
Baboon	7.574	7.9974
Barbara	7.509	7.9971
Peppers	7.329	7.9972

7.3. UACI & NPCR TESTS

UACI tests determine the difference between the original and encrypted images. The variation rate between image pixels before and after the encryption is determined using a NPCR test. Equations 6, 7, and 8 are used to measure the NPCR and UACI tests [29].

$$UACI = \frac{1}{RC}\left[\sum_{i,j}\frac{|P_1(i,j)-P_2(i,j)|}{255}\right] * 100 \quad (6)$$

$$NPCR = \frac{\sum_{i,j} N(i,j)}{R*C} * 100 \quad (7)$$

$$N(i,j) = \begin{cases} 0, & if\ P_1(i,j) = P_2\ (i,j) \\ 1, & if\ \ P_1(i,j)\ \neq P_2\ (i,j) \end{cases} \quad (8)$$

Where P2 (i, j) is the original image, P1 (i, j) is an encrypted image, C is the image's width, and R is the

height of the image. Table 7 shows the results of applying the tests (NPCR and UACI) to the images used in the experiments after applying the proposed method.

Table 7. NPCR and UACI values for the experiment image

Image	NPCR			UACI		
	Red	Green	Blue	Red	Green	Blue
Lena	99.63	99.65	99.63	33	32	33
Baboon	99.61	99.60	99.62	30	28	31
Barbara	99.61	99.64	99.63	29	29	31
Peppers	99.62	99.60	99.59	29	34	34

7.4. CORRELATION TEST

The correlation test is one of the most important statistical tests used to assess the strength and effectiveness of the encryption system. An encryption algorithm removes and destroys the strong correlation between each pixel and its neighbors in the original image [30]. Use Equations 9, 10, and 11 to calculate the correlation coefficient.

$$r_{p,q} = \frac{\sum_{i=1}^{M}(p_i-\bar{p})(q_i-\bar{q})}{\sqrt{[\sum_{i=1}^{M}(p_i-\bar{p})^2][\sum_{i=1}^{M}(q_i-\bar{q})^2]}} \quad (9)$$

$$\bar{p} = \frac{1}{M}\sum_{i=1}^{M} p_i \quad (10)$$

$$\bar{q} = \frac{1}{M}\sum_{i=1}^{M} q_i \quad (11)$$

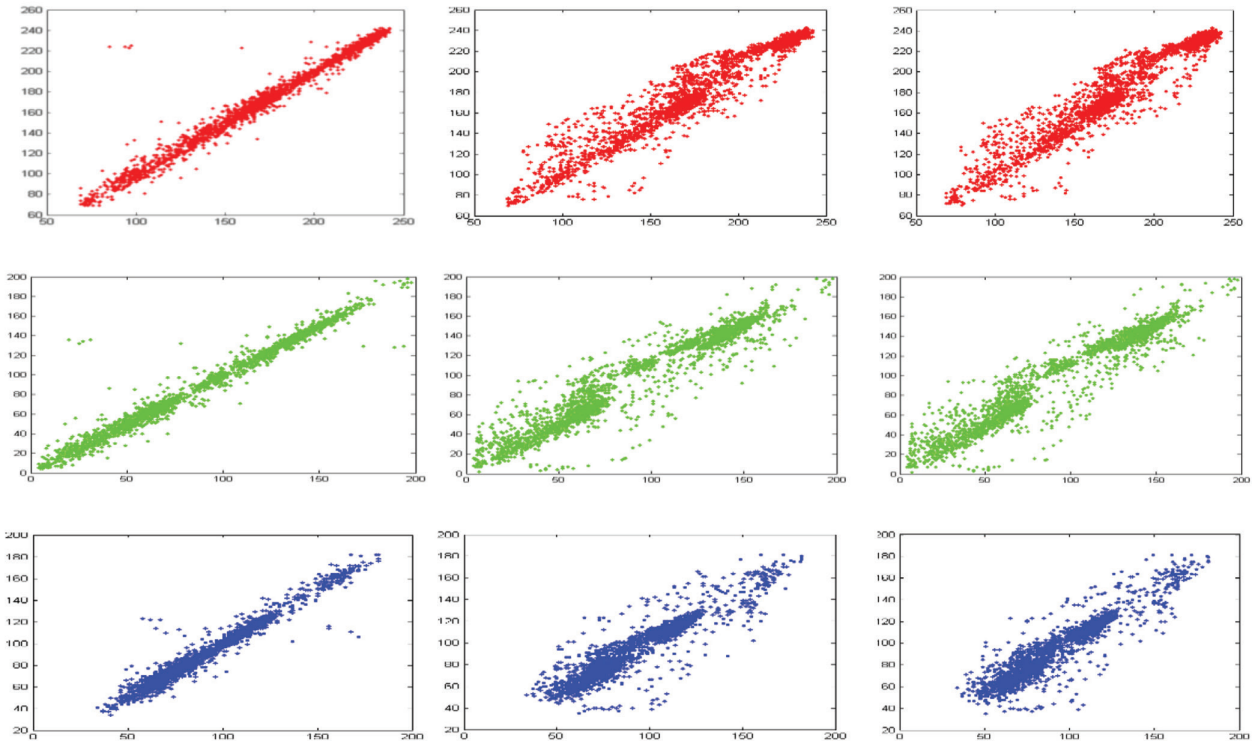
Where (pi) represents the value of the pixel (i-th) selected, (qi) represents the value of the adjacent pixel, and (M) represents the number of pixels selected from the image. Fig. 5 shows the results of applying the correlation test to the Lena image in the horizontal, vertical, and diagonal directions before and after applying the suggested method for 3000 pixels. Tables 8 and 9 show the correlation coefficient values for the test images before and after the suggested method was applied. Original image correlation coefficients are near 1, while cipher image correlation coefficients are close to 0. The results show that the proposed method's correlation coefficients achieve outstanding results close to 0.

Table 8. Correlation coefficients values for the original images

Image	Horizontal	Vertical	Diagonal
Lena	0.9534	0.9866	0.9540
Baboon	0.7289	0.6122	0.6324
Barbara	0.8974	0.9123	0.8427
Peppers	0.8780	0.9460	0.8542

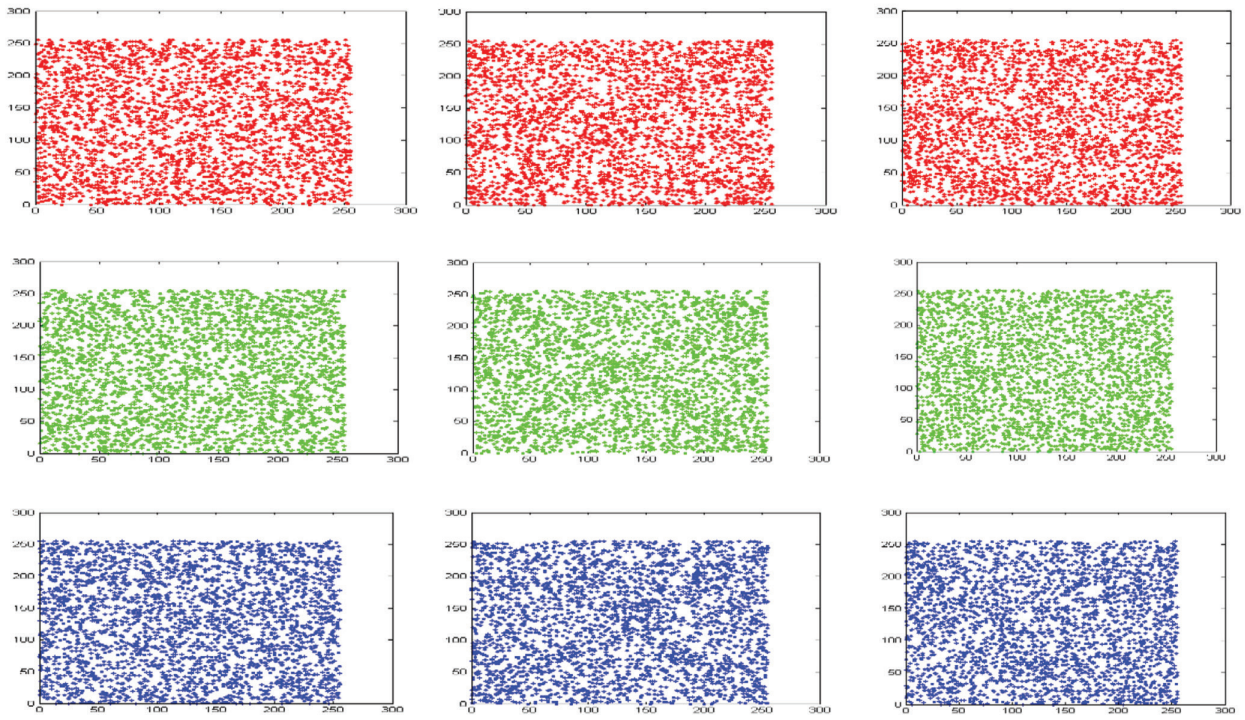
Table 9. Correlation coefficient values for the encrypted images

Image	Horizontal	Vertical	Diagonal
Lena	0.0096	-0.0071	-0.0079
Baboon	0.0016	-0.0023	-0.0087
Barbara	0.0030	-0.0068	-0.0147
Peppers	-0.0008	0.0015	-0.0134



Vertical Horizontal Diagonal

(a)



Vertical Horizontal Diagonal

(b)

Fig. 5. The RGB correlation test for the Lena's image (a) Original image (b) Encrypted image

7.5. PSNR ANALYSIS

One of the most crucial tests to evaluate the effectiveness of image coding systems is PSNR. When the PSNR value is less than 10, the encryption system's strength is excellent and ideal [31]. MSE represents the difference between the encrypted image and the original, and the optimal value of MSE is when it is very high [30], [31]. Use Equations 12 and 13 to calculate the PSNR and MSE values.

$$MSE = \frac{1}{mn}\sum_{ij}(M(i,j) - N(i,j))^2 \quad (12)$$

$$PSNR = 10log_{10}\left(\frac{P^2}{MSE}\right) \quad (13)$$

Where M (i, j) is an encrypted image, N (i, j) is the original image, (i, j) is the coordinate, and P is the range of pixels in the image [32]. Table 10 shows the results of MSE and PSNR values on the images after applying the proposed method. The results show differences between the original and encrypted image, as explained by the high MSE and low PSNR values.

Table 10. PSNR and MSE for the encrypted images

image	PSNR			MSE		
	Red	Green	Blue	Red	Green	Blue
Lena	7.845	8.558	9.548	106.80	90.63	72.15
Baboon	8.869	9.431	8.518	84.35	74.11	91.46
Barbara	9.016	8.956	8.444	81.55	82.67	93.03
Peppers	9.218	7.692	7.524	77.85	110.63	114.97

We compare the results of applying the proposed method with the results of other methods applied to the Lena image to determine the strength and effectiveness of the proposed method. Table 11 presents a comparison of the test results for entropy, MSE, PSNR, and correlation using the suggested method and other researchers' results on the Lena image.

Table 11. A comparison between the proposed method and some other methods.

Method	Entropy	NPCR	UACI	Correlation		
				H	V	D
Proposed method	7.997	99.64	32.66	0.0096	-0.0071	-0.0079
REF. [4]	7.997	99.59	33.28	0.0016	-0.0020	0.0047
REF. [15]	7.999	99.61	33.47	-0.0001	0.0002	-0.0001
REF. [16]	7.997	99.62	33.42	0.0058	0.0033	0.0010
REF. [18]	7.997	99.63	33.37	-0.0010	-0.0030	-0.0051

8. CONCLUSION

Encryption is one of the most important ways to maintain the secrecy of data and prevent unauthorized access. This paper suggests a new encryption method for data at two levels to increase data security. The first level of protection combines LFSR and DNA coding, while the second uses chaotic maps. The randomness of LFSR in rearranging the locations of pixels at the first level and the 3D chaotic map in data encryption at the second give strength and robustness to this method. Several statistical tests were carried out to prove the ef-

fectiveness of the proposed method. The results of the tests revealed a high level of security compared to the results of other methods. Future work should aim to use other types of chaotic maps and DNA coding.

ACKNOWLEDGMENTS

The authors would like to thank Mustansiriyah University (www.uomustansiriyah.edu.iq) in Baghdad, Iraq, for supporting this work. The authors also appreciate the comments of reviewers.

9. REFERENCES

[1] Z. Tang, Z. Yin, R. Wang, X. Wang, J. Yang, J. Cui, "A Double-Layer Image Encryption Scheme Based on Chaotic Maps and DNA Strand Displacement", Journal of Chemistry, Vol. 2022, 2022.

[2] J. Lin, K. Zhao, X. Cai, D. Li, Z. Wang, "An Image Encryption Method Based on Logistic Chaotic Mapping and DNA Coding", Proceedings of Remote Sensing Image Processing, Geographic Information Systems, and Other Applications, Vol. 11432, 2019.

[3] J. Chauhan and A. Jain, "Survey On Encryption Algorithm Based On Chaos Theory And DNA Cryptography", International Journal of Advanced Research in Computer and Communication Engineering, Vol. 3, No. 8, 2014, pp. 7801-7803.

[4] P. N. Lone, D. Singh, U. H. Mir, "Image encryption using DNA coding and three-dimensional chaotic systems", Multimedia Tools and Applications, Vol. 81, No. 4, 2022, pp. 5669-5693.

[5] S. T. Allawi, "Image Encryption Based on Chaotic Mapping and Random Numbers", Journal of Engineering and Applied Sciences, Vol. 14, No. 19, 2019, pp. 6954-6958.

[6] Z. Li, C. Peng, W. Tan, L. Li, "A Novel Chaos-Based Image Encryption Scheme by Using Randomly DNA Encode and Plaintext Related Permutation", Applied Sciences., Vol. 10, No. 21, 2020, pp. 1-19.

[7] Q. S. Alsaffar, H. N. Mohaisen, F. N. Almashhdini, "An encryption based on DNA and AES algorithms for hiding a compressed text in colored Image", IOP Conference Series: Materials Science and Engineering, Vol. 1058, No. 1, 2021, p. 012048.

[8] Y. Wang, X. Li, Q. Wang, " Integral Imaging Based Optical ImageEncryption Using CA-DNA Algorithm", IEEE Photonics Journal, Vol. 13, No. 2, 2021.

[9] T. V. Medha Sreenivasan, A. Sidhardhan, V. M. Priya, "5D Combined Chaotic System for Image Encryption with DNA Encoding and Scrambling", Proceedings of the International Conference on Vision Towards Emerging Trends in Communication and Networking, Vellore, India, 30-31 March 2019.

[10] N. Iqbal, S. Abbas, M. A. Khan, T. Alyas, A. Fatima, A. Ahmad, "An RGB Image Cipher Using Chaotic Systems, 15-Puzzle Problem and DNA Computing", IEEE Access, Vol. 7, 2019, pp. 174051-174071.

[11] P. Vinotha, D. Jose, "VLSI Implementation of Image Encryption Using DNA Cryptography", Intelligent Communication Technologies and Virtual Mobile Networks, Springer Nature, Switzerland, 2020, pp. 190-198.

[12] K. S. Kumari and C. Nagaraju, "DNA encrypting rules with Chaotic Maps for Medical Image Encryption", Proceedings of the 5th International Conference on Intelligent Computing and Control Systems, Madurai, India, 6-8 May 2021 pp. 832-837.

[13] Nalini M. K and Radhika K. R, "Secured Key Generation for Biometric Encryption using Hyper-Chaotic Map and DNA Sequences", SSRN Electronic Journal, 2021, pp. 585-595.

[14] A. Pai, P. K. Pareek, G. Prasad, P. Singh, B. K. Deshpande, "Image Encryption Method by Using Chaotic Map and DNA Encoding", Natural Volatiles & Essential Oils, Vol. 8, No. 5, 2021, pp. 10391-10400.

[15] S. M. Hameed, I. A. Taqi, "A New Beta Chaotic Map with DNA Encoding for Color Image Encryption", Iraqi Journal of Science, Vol. 61, No. 9, 2020, pp. 2371-2384.

[16] K. H. Moussa, H. G. Mohamed, D. H. ElKamchouchi, "A Novel Color Image Encryption Algorithm Based on Hyperchaotic Maps and Mitochondrial DNA Sequences", Entropy, Vol. 22, No. 158, 2020, pp. 1-15.

[17] J. Zheng and L. F. Liu, "Novel Image Encryption by Combining Dynamic DNA Sequence Encryption and the Improved 2D Logistic Sine Map," IET Image Process., vol. 14, no. 11, pp. 2310–2320, 2020, doi: 10.1049/iet-ipr.2019.1340.

[18] A. Girdhar, V. Kumar, "A RGB image encryption technique using Lorenz and Rossler chaotic system on

DNA sequences", Multimedia Tools and Applications, Vol. 77, No. 20, 2018, pp. 27017-27039.

[19] M. S. Fadhil, A. K. Farhan, M. N. Fadhil, "Designing Substitution Box Based on the 1D Logistic Map Chaotic System", IOP Conference Series: Materials Science and Engineering, Vol. 1076, No. 1, 2021, p. 012041.

[20] R. Ismail Abdelfattah, H. Mohamed, M. E. Nasr, "Secure Image Encryption Scheme Based on DNA and New Multi Chaotic Map", Journal of Physics: Conference Series - IOPscience, Vol. 1447, No. 1, 2020.

[21] C. L. Chunhu Li, G. Luo, "An Image Encryption Scheme Based on The Three-Dimensional Chaotic Logistic Map", International Journal of Network Security, Vol. 21, No. 1, 2019, pp. 22-29.

[22] M. B. Hossain, M. T. Rahman, A. B. M. S. Rahman, S. Islam, "A New Approach of Image Encryption Using 3D Chaotic Map to Enhance Security of Multimedia Component", Proceedings of the International Conference on Informatics, Electronics & Vision, Dhaka, Bangladesh, 23-24 May 2014.

[23] K. Singh, K. Kaur, "Image Encryption using Chaotic Maps and DNA Addition Operation and Noise Effects on it", International Journal of Computer Applications, Vol. 23, No. 6, 2011, pp. 17-24.

[24] K. A. Kumari, B. Akshaya, B. Umamaheswari, K. Thenmozhi, R. Amirtharajan, P. Praveenkumar, "3D Lorenz Map Governs DNA Rule in Encrypting DICOM Images", Biomedical and Pharmacology Journal, Vol. 11, No. 2, 2018, pp. 897-906.

[25] B. Wang, S. Zhou, X. Zheng, C. Zhou, J. Dong, L. Zhao, "Image watermarking using chaotic map and DNA coding", Optik, Vol. 126, No. 24, 2015, pp. 4846-4851.

[26] S. J. Sheela, K. V. Suresh, D. Tandur, "A Novel Audio Cryptosystem Using Chaotic Maps and DNA Encoding", Journal of Computer Networks and Communications, Vol. 2017, 2017.

[27] F. A. Salman, K. A. Salman, "Enhanced Image Encryption Using Two Chaotic Maps", Journal of ICT Research and Applications, Vol. 14, No. 2, 2020, pp. 134-148.

[28] M. Tanveer et al. "Multi-Images Encryption Scheme Based on 3D Chaotic Map and Substitution Box", IEEE Access, Vol. 9, 2021, pp. 73924-73937.

[29] S. Agarwal, "Secure Image Transmission Using Fractal and 2D-Chaotic Map", Journal of Imaging, Vol. 4, No. 1, 2018.

[30] A. M. Alabaichi, "Color Image Encryption using 3D Chaotic Map with AES Key Dependent S-Box", International Journal of Computer Science and Network Security, Vol. 16, No. 10, 2016, pp. 105-115.

[31] C. Liu, Q. Ding, "A Color Image Encryption Scheme Based on a Novel 3D Chaotic Mapping", Complexity, Vol. 2020, 2020.

[32] S. T. Allawi, M. M. Abbas, R, H. Mahdi, "New Method For Using Chaotic Maps To Image Enceyption", International Journal of Civil Engineering and Technology, Vol. 9, No. 13, 2018, pp. 244-231.

Smart Bagged Tree-based Classifier optimized by Random Forests (SBT-RF) to Classify Brain-Machine Interface Data

Original Scientific Paper

Omar A. Sesa
Mansoura University,
Computers and Control Systems Engineering Dept.,
Faculty of Engineering,
Mansoura, Dakahlia, Egypt
omarsesa.os@std.mans.edu.eg

Amira Y. Haikal
Mansoura University,
Computers and Control Systems Engineering Dept.,
Faculty of Engineering,
Mansoura, Dakahlia, Egypt
amirayh@mans.edu.eg

Mostafa A. Elhosseini
Mansoura University,
Computers and Control Systems Engineering Dept.,
Faculty of Engineering, Mansoura, Dakahlia, Egypt
melhosseini@mans.edu.eg
Taibah University,
College of Computer Science and Engineering in Yanbu, Yanbu, Madinah, Saudi Arabia
melhosseini@ieee.org

Hesham H. Gad
Mansoura University,
Computers and Control Systems Engineering Dept.,
Faculty of Engineering, Mansoura, Dakahlia, Egypt
hmgad2004@mans.edu.eg

Abstract *– Brain-Computer Interface (BCI) is a new technology that uses electrodes and sensors to connect machines and computers with the human brain to improve a person's mental performance. Also, human intentions and thoughts are analyzed and recognized using BCI, which is then translated into Electroencephalogram (EEG) signals. However, certain brain signals may contain redundant information, making classification ineffective. Therefore, relevant characteristics are essential for enhancing classification performance. . Thus, feature selection has been employed to eliminate redundant data before sorting to reduce computation time. BCI Competition III Dataset Iva was used to investigate the efficacy of the proposed system. A Smart Bagged Tree-based Classifier (SBT-RF) technique is presented to determine the importance of the features for selecting and classifying the data. As a result, SBT-RF is better at improving the mean accuracy of the dataset. It also decreases computation cost and training time and increases prediction speed. Furthermore, fewer features mean fewer electrodes, thus lowering the risk of damage to the brain. The proposed algorithm has the greatest average accuracy of ~98% compared to other relevant algorithms in the literature. SBT-RF is compared to state-of-the-art algorithms based on the following performance metrics: Confusion Matrix, ROC-AUC, F1-Score, Training Time, Prediction speed, and Accuracy.*

Keywords: *Brain-Machine Interface; Bagged Trees; Classification; Feature Selection; Optimization; Random Forests*

1. INTRODUCTION

A BCI uses electrodes or sensors to interface machines with the human brain [1] based on neurosciences [2]. BCI receives and transmits electrical signals, which can help doctors to discover more information about brain issues and diseases like stroke to use in rehabilitation [3], [4]. It can also simulate a human brain to enhance machine learning and control objects as natural parts of its body representation[5]–[7]. BCI helps the medical field & health care and plays a crucial role in several areas, such as entertainment, education, marketing, and automated control [8]–[10]. Freely available datasets relevant to BCI can be used to test the suggested technique.

One of these datasets is BCI competitions. The Brain-Computer Interface (BCI) Competition was established to evaluate signal processing and classification methods for BCIs. BCI Competition III dataset IVa, five healthy subjects' brains activated for some MI tasks, namely, "aa," "al," "av," "aw," and "ay," were recorded [11]. This recording was taken using Brain Amp amplifiers and a 128-channel Ag/AgCl electrode cap from Electrocerebral Inactivity (ECI) and 118 channels in those five healthy subjects by putting 118 electrodes international 10/20 system [12].

The feature selection algorithm is an important preprocessing mathematical stage to decrease data size by

reducing the dimension of the data set by removing irrelevant and unnecessary data and selecting the most suitable properties for the data categories. In addition, feature selection increases speed and accuracy, thus obtaining high performance [13], [14]. Feature selection methods can be widely categorized into filter, wrapper, and embedded methods[15], [16]. For example, RF is classified as embedded feature selection because it combines wrapper and filters' positive aspects such as high speed and high accuracies [17] proposed by Breiman [18].

Data is typically classified into a class or category via classification algorithms. There are three types of classification: binary classification, multiclass classification, and multilabel classification. Binary classification techniques are used to classify datasets with only two classes: normal and abnormal states, commonly referred to as "class 0" and "class 1," respectively. Datasets with more than two classes are classified using multiclass classification techniques. Some binary classification techniques can also be utilized for multiclass classification. Finally, multilabel classification techniques classify datasets with two or more classes, with one or more class labels predicted for each input [19], [20].

The selected features are classified using the Bagged Trees algorithm (BT). Researchers bolstered weak classification accuracies with a combination of classifiers, like bagging or kernel additions. A bagging algorithm combines multiple classifiers (Ensemble) proposed by Breiman. Bagging is a way to make the poor classifier better than the first by sampling the dataset into bootstrap samples to prevent the classifier from getting overfitted [21], [22] by learning every classifier individually and collecting the votes of the classifiers. The class with the highest number of votes is the winner. Bagging the decision tree will produce the BT algorithm [23]–[25].

Based on the details listed above, SBT–RF algorithm has been built from the Random Forests (RF) algorithm to choose dataset features and then apply the BT algorithm to classify the selected features. As a result, the proposed algorithm outperforms all different algorithms in accuracy testing.

The main contributions of this paper are:

- Delete superfluous channels by using RF to select features.
- Bagging the decision trees produces the BT classifier that will classify the dataset.
- Reducing the computational cost and time required to train and predict BCI datasets classes.

This paper is structured as follows: Section 2 covers the relevant work and emphasizes the pros and cons of each given input, Section 3 presents the main intended achievement of solving the previously mentioned problems through SBT-RF, Section 4 consists of dataset description, the used performance metrics and the machine results obtained. Finally, section 5 discusses the conclusion.

2. RELATED WORK

Literature on machine learning and feature selection application to BCI models has increased in recent years. For example, Md.A.M. Joadder et al. (2019) [26] developed a method to classify mental states for an SI-based BCI system. First, they applied several feature extraction techniques; Katz Fractal Dimension (KFD), Sub-band Energy, Log Variance, and Root Mean Square (RMS). After that, they used the obtained features as input to Linear Discriminant Analysis (LDA) classifier, and the best average accuracy value is 84.35% by KFD with LDA. However, this algorithm needs a large number of channels to classify well.

Yongkoo Park et al. (2019) [27] proposed a new method; Local Region Frequency Optimized Common Spatial Pattern (LRFCSP). The features are extracted from the best local regions by applying Variance Ratio Dispersion Score (VRDS) and Interclass Feature Distance (ICFD) techniques to optimize CSPs. Unfortunately, with a mean classification accuracy of 92.93%, The approach is ineffective when classifying tiny samples.

Amin Hekmatmanesh et al. (2020) [28] proposed a technique to enhance a common spatial pattern algorithm to recognize and classify BCI Competition III dataset IVa by combining four different algorithms. Those were Kernel Linear Discriminant Analysis (KLDA), the Kernel Principal Component Analysis (KPCA), the Soft Margin Support Vector Machine (SSVM) classifier, and the Generalized Radial Bases Functions (GRBF) to create methods called DFBCSP DSLVQ SSVM GRBF with an average accuracy of 92.70%. Still, this method increases the error ratio for multiclass.

Sahar Selim et al. (2020) [29] compare feature recognition techniques; RMS, Renyi entropy, Shannon entropy, Katz fractal dimension, and CSP to minimize the number of features used as input to the classifier. The classifiers they used are Support Vector Machine (SVM) and Linear Discriminant Analysis (LDA) as a classifier, and the best average accuracy among all those algorithms is 79.46%. Furthermore, Renyi entropy with LDA and SVM achieved reasonable accuracy with high computational speed, which executes faster than CSP in computational time but with lower precision.

Yao Guo et al. (2020) [4] proposed two methods. The first is the Filtered Band Component Regularized Common Spatial pattern (FCCSP) to increase the robustness of CSP against the small samples by reducing estimations and adding priorities to the spatial filter. The second is Minimum Redundancy Maximum Relevance (mRMR) to discard unnecessary features and use the rest as input to LDA Classifier, and the highest average accuracy obtained was by FCCSP (82.01%). However, this approach may not perform well with multiclass motor imagery tasks.

Wenlong Hang et al. (2020) [30] proposed a new framework that includes a Support Matrix Machine

(SMM) as a core part of a deep stacked network to build a Deep Stack Support Matrix Machine (DSSMM). The framework has been applied to BCI competition III dataset IVa and performs effectively with feed-forward with an average accuracy of 83.68%. On the other hand, executing this method takes a high computational time.

Research gap: In light of the studies mentioned above, it appears that most of these methods suffer from the high amount of time spent versus poor quality in large or small sample settings. Furthermore, they have relatively high error margins when attempting to classify datasets. Table 1 summarizes the pros and cons of these methods.

Table 1. Summaries previously discussed algorithms focusing on various pros and cons

Author	Method	Mean Accuracy	No. of Channels	Pros	Cons
Md.A.M. Joadder et al. [26]	Katz + LDA	84.35%	118	-	This algorithm needs a large number of channels to classify well.
Yongkoo Park et al. [27]	LRFCSP	(92.93±3.99) %	18	Performs well with large sets	This method obtains poor performance when classifying a small sample set.
Amin Hekmatmanesh et al. [28]	DFBCSP DSLVQ SSVM GRBF	92.70%	118	Performs well with binary classes	This method increases the error ratio for multiclass.
Sahar Selim et al. [29]	CSP+LDA	79.77%	18	Executes fast	Low on accuracy.
Yao Guo et al. [4]	FCCSP	82.01%	118	Performs well with binary classes	This approach may not perform well with multiclass motor imagery tasks.
Wenlong Hang et al. [30]	DSSMM	83.68%	118	-	High computational time

3. MATERIALS AND METHODS

There are four components to a BCI: Data collection, Preprocessing, Feature selection, and classification.

3.1. DATA COLLECTION

The BCI Competition III Dataset Iva and other biomedical datasets have been used to evaluate the efficiency of the proposed approach.

3.1.1. BCI Dataset

This BCI Competition III Dataset IVa contains Motor imagery (MI) tasks of five healthy subjects, namely, "aa," "al," "av," "aw," and "ay" [31], which are recorded. The EEG data consists of three classes (right hand(R), Left hand (L), and foot (F)) [32]. Only cues for the classes "right hand" and "foot" were provided. There were two types of visual stimulation. The first type, where targets were indicated by letters appearing behind a fixation cross (which might motivate a little bit of target-correlated eye movements). The second and (2) where a randomly moving object indicated targets (inducing target-uncorrelated eye movements). This dataset was acquired from the participants seated on a chair well relaxed. Visual cues were displayed for 3.5 s, during which the subject had to perform the MI tasks; left or right hand and foot [33], as shown in (Fig. 1).

Brain Amp amplifiers, an ECI 128-channel Ag/AgCl electrode cap, and 118 electrodes were used [34] as defined by the international 10/20 system, as shown in (Fig. 2). The BCI Competition III Dataset Iva and other biomedical datasets have been used to evaluate the efficiency of the proposed approach.

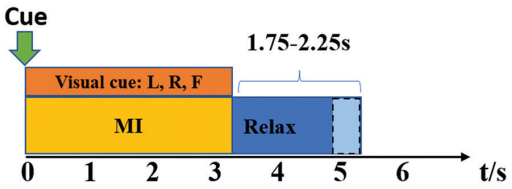


Fig 1. Procedure timeline.

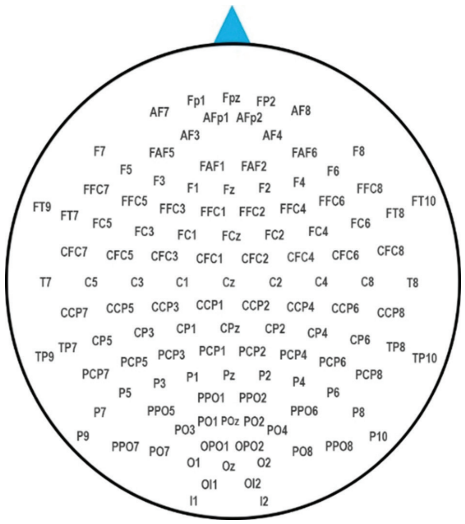


Fig. 2. 118-channel positions

The EEG segments representing the MI part were separated from the dataset. For that reason, the marker's position that indicated the start of 280 cues and the actuality that each movement was 3.5 s long had been utilized. Thus, 280 EEG segments were obtained for each subject toward the finish of this procedure, as described in Table 2.

Table 2. Dataset trials.

Among 280 trials					
Subject	**aa**	**al**	**av**	**aw**	**ay**
Size of data with two classes (RH and RF)	298458 × 118	283574 × 118	283042 × 118	282838 × 118	283562 × 118
Number of trials considered as a training trial with the class label	168	224	84	56	28
Number of trials considered as a testing trial without a class label	112	56	196	224	252

The EEG segments representing the MI part were separated from the dataset. For that reason, the marker's position that indicated the start of 280 cues and the actuality that each movement was 3.5 s long had been utilized. Thus, 280 EEG segments were obtained for each subject toward the finish of this procedure, as described in Table 2.

3.2. SIGNAL PREPROCESSING

Raw EEG data are cluttered with noise and artifacts. Signals are preprocessed to remove artifacts. Among the functions are artifact rejection, channel selection, and baseline filtering. A Butterworth bandpass filter of the fifth order can be used in this case. Preprocessing of the signals is not our primary objective in this paper.

3.3. FEATURE SELECTION

Choosing features is a critical part of designing a machine-learning model. As well as helping to distinguish relevant from irrelevant attributes, it decreases the dimensionality of the original dataset, which helps improve performance. Furthermore, the selection of relevant attributes helps increase learning performance. The application of feature selection allows models to be interpreted after a short training period, which is essential for improving brain-computer interfaces based on motor imagery [35]. The RF explains the significance of each feature [18].

Feature selection using the RF algorithm has two targets [36], [37]–[41] first to discover highly related feature variables and find features with comparatively little data and better capacity to communicate anticipating results. An evaluation feature is usually valued according to two measures: the Gini index, also known as the Gini coefficient or the Gini impurity. It indicates the likelihood that a variable will be wrongly classified when selected at random. The second measure is the error rates of OOB, as shown in Algorithm 1. If z is the total number of features $\{x_1, x_2, \ldots\ldots x_z\}$ then the particular procedures to calculate feature importance for every x_j as follows:

- Applying the bootstrap method to the original training dataset creates stochastic sample sets K by fetching them for classification or regression. Then, send them back to the original dataset to create another stochastic sample set, and every time this procedure happens, it makes an OOB sample for the unsampled data.
- m_{try} features ($m_{try} \leq n$) are haphazardly extricated at every node of every tree in the RF algorithm as a randomly produced feature subset by ascertaining the data contained in each feature and calculating it. Feature with the best possible classification capability is chosen among the m_{try} features to split the node, which makes decision trees more diverse.
- Feature importance can be calculated by the Gini index as follows: For each feature x_j The Gini index value can be calculated by getting the average of the Gini index change amount before and after node impurity gets split in whole decision trees of the RF algorithm as shown in (Fig. 3). The Gini index can be calculated as follows [42],[43], [38]:

$$GI_m = \sum_{k=1}^{|k|} \sum_{k' \neq 1} P_{mk} P_{mk'} = 1 - \sum_{k=1}^{|k|} P_{mk}^2 \quad (1)$$

Where GI_m refers to the Gini index of node m, and K refers to K-categories, and P_{mk} refers to the category k proportion in node m. x_j the amount of Gini index change calculates the feature importance value at node m before and after node m gets split.

$$MVI_{jm}^{(Gini)} = GI_m - GI_l - GI_r \quad (2)$$

MVI represents the variable importance. GI_lrepresents the Gini index of the left branch and GI_r represent the Gini index also but for the right branch.

Feature x_j in decision tree I and belongs to set M, the importance of this feature [18],[44]:

$$MVI_{ij}^{(Gini)} = \sum_{m \in M} MVI_{jm}^{(Gini)} \quad (3)$$

Let N be the number of trees that are created in RF, then:

$$MVI_j^{(Gini)} = \sum_{i=1}^{N} MVI_{ij}^{(Gini)} \quad (4)$$

Normalize the whole scores of feature importance by:

$$MVI_j = \frac{MVI_j^{(Gini)}}{\sum_{j=1}^{C} MVI_j^{(Gini)}} \quad (5)$$

Algorithm 1: Random Forests

1 **for** features x_j, j = 1 to z do
2 **for each** tree learner in sample k = 1 to |k| do
3 Find all nodes m that use x_j.

4 | | Obtain feature importance by calculating impurity before and after splitting of x_j using the Gini index.

5 | | Aggregate the enhancements

6 | **end**

7 | Aggregate the improvements of all the decision trees to get feature importance of x_j.

8 **end**

---

1 **while** developing tree

2 | Calculate accuracy for OOB observations and record the prediction

3 **end**

4 Do premutation to x_j 00B observations to breakdown the relations between the target and the permuted feature x_j.

5 **for each** OOB observation

6 | Calculate the accuracy and record the prediction again

7 **end**

8 Get the average of the performance reduction that had been happened due to permute xj and utilize it to measure the xj feature importance.

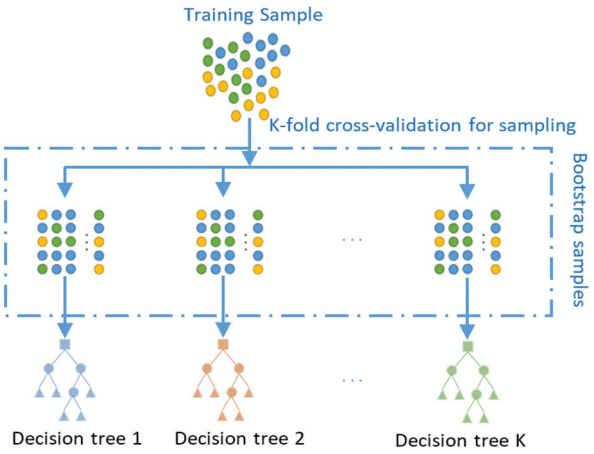


Fig. 3. RF algorithm process

3.4. CLASSIFICATION

A detailed explanation of the main algorithms is provided in this section.

3.4.1. Cart Decision Tree

The CART algorithm has been proposed by Breiman [42] to create generations of decision trees. The CART is a decision tree depending on binary methodology to generate a decision tree from the dataset [45]–[47], as shown in (Fig. 4). Using the CART as a classifier depends on the Gini index. The Gini index is applied on each node of the tree to detect the effectiveness of the data attribute to choose where the effective split will be in the set to provide new subsets [21], [47], as presented in Algorithm 2. The smaller the Gini value of the attribute, the greater the purity of the node, and it will be chosen as the best one to do the splitting. The algorithm is represented mathematically [43] as follows:

D refers to a dataset, kth proportion sample of D

P_k (k=1,2…..|k|)

Gini(D) shows the probability between 2 classes that have been randomly chosen from D

$$\text{Gini(D)} \propto \frac{1}{\text{the purity of D}} \tag{6}$$

Smaller the Gini value, the higher the purity. The Gini index of an attribute x can be obtained as follow:

$$\text{Gini\_index(D, x)} = \sum_{v=1}^{v} \frac{|D^v|}{|D|} Gini(D^v) \tag{7}$$

Where D^v is a subset of D with tuples having value (v) For example, $x_* = arg\ min_{x \in A}\ Gini\_index(D,x)$.Where A: the candidate attribute set.

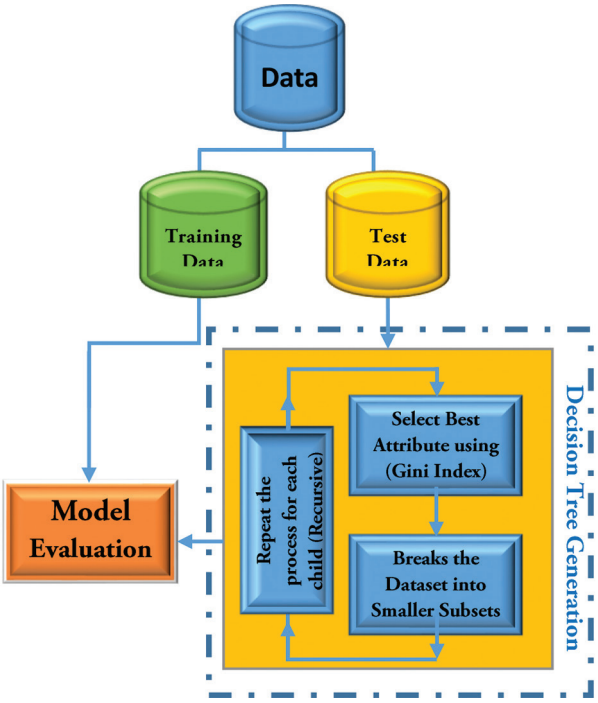


Fig. 4. CART decision trees

Algorithm 2: CART

---

1 **if** Number of Feature values=k Then

2 | Possible splits =k-1

3 **end**

4 **for each** feature(F) in Dataset(D)

5 | **while** Split No. (SN)<k

6 | | Find the best split (BS) using Gini Index

7 | **end**

8 | Split the feature at BS

```
while the Stopping criterion is not satisfied
    for each node(N)
        Find the node's best split (NBS) using Gini Index
    end
    Split the node at NBS
end
end
```

3.4.2. Classification Algorithm

As Breiman [22] proposed, bagging is an algorithm that produces multiple versions of a predictor that can be aggregated. For example, aggregation averaging the outcome by bagging a regressor algorithm to predict a numerical value [48], [49].

Bagging a classifier algorithm is done by collecting the number of votes for each classifier version to predict the output class. Algorithm 3 shows how the bagging algorithm works in steps on the dataset. First, the multiple versions are created by making frequent bootstraps of the learning set, which will be used as new learning sets [50], [51], as shown in (Fig. 5).

In classification [22], [52] a predictor $\varphi(x,L)$ predicts class label $j \in \{1,\ldots,j\}$

$$Q(j|x) = P(\phi(x,\mathcal{L}) = j) \tag{8}$$

The meaning of $Q(j|x)$ is this: over many independent replicates of the learning set L, ϕ predicts class label j at input x with relative frequency $Q(j|x)$ Let $P(j|x)$ be the probability that input x generates class j. Then the likelihood that the predictor classifies the developed state at x correctly is

$$\sum_j Q(j|x)P(j|x) \tag{9}$$

The total probability of the right classification is

$$r = \int \left[\sum_j Q(j|x)P(j|x)\right] Px(dx) \tag{10}$$

where $Px(dx)$ is the probability distribution of x

Please note this for any $Q(j|x)$

$$\sum_j Q(j|x)P(j|x) \le \max_j P(j|x) \tag{11}$$

with equality only if

$$Q(j|x)\begin{cases} 1 \; if \; P(j|x)\, max_i = P(i|x) \\ 0 \; else \end{cases} \tag{12}$$

The predictor $\phi^*(x)=argmax_i P(j|x)$ (defined as the Bayes predictor) conduces to the previous representation for $Q(j|x)$ and achieves the highest possible correct classification rate:

$$r^* = \int \max_j P(j|x)Px(x) \tag{13}$$

Call ϕ order correct at the input x if

$$argmax_j Q(j|x) = argmax_j P(j|x) \tag{14}$$

It implies that if input x occurs more frequently than any other variable in class j, then ϕ also predicts class j at x more regularly than the others. An order-correct predictor is not always an accurate predictor. So, the aggregated predictor is:

$$\phi_A(x) = \; argmax_j Q(j|x) \tag{15}$$

The correct classification probability at x for the aggregated predictor is:

$$\sum_j I(argmax_i Q(i|x) = j)\; P(j|x) \tag{16}$$

$I(.)$ refer to the indicator function. If ϕ is ordered correct at x, then the previous equation equals$max_j P(j|x)$.

Let us assume that C is the set of all inputs x at which ϕ is correctly ordered, the expression for the proper classification probability of ϕ_A will be:

$$r_A = \int_{x\in C} \max_j P(j|x)Px(dx) + \int_{x\in C'} \left[\sum_j I(\phi_A(x) = j)P(j|x)\right] Px(x) \tag{17}$$

According to the previous equations, if the predictor has a good sensation to predict order correct for most inputs of x, then aggregation can convert it into an almost optimum predictor. In contrast to the numerical prediction situation, weak predictors can be converted into bad ones. Bagging unstable classifiers improve them [22].

Algorithm 3: Bagging

```
Variables: OD: Original Dataset
           N: Number of bootstrap samples
           L: Learning Algorithm
           C*: Bagging ensemble classifier
for j=1 to N
    BSSj ← bootstrap sample from OD
    Create Classifier Cj ←L(BSSj)
end
for each new instance, predict the class label
    C*(x) = arg max_y Σ_{j=1}^{N} [C_j(x) = y]
end
```

In machine learning, Cross-validation is a blind technique used frequently to enhance model prediction and reduce bias. The BCI dataset has been divided randomly into k sets for each subject individually (k-fold cross-validation) [53]. In this study, 5-fold cross-validation has been used. One of these sets was utilized as a testing set, while the other four were used as training sets. This method is repeated five times, each time with a different set.

Previous work has suffered from weaknesses since it was designed as an independent algorithm. However, such defects can be overcome by combining different algorithms. Therefore, this work combines BT and RF algorithms to perform better than operating individually.

The BT algorithm reduces the pre-steps needed to obtain high-dimensional feature data, including dimension reduction and feature selections. Thus, from the speed evaluation perspective, the BT algorithm can speed up the process impressively while maintaining the easiness of creating parallel methods and the simplicity of implementation. Furthermore, if things go wrong, it can fix the error, as balancing errors occur by unbalanced data sets or trying to reduce the impact of losing a huge part of features by maintaining accuracy despite the current condition.

On the other hand, the RF algorithm focus on getting the best feature selection possible by calculating the Gini index and OOB error, which leads to judging the importance of features and interaction among different ones.

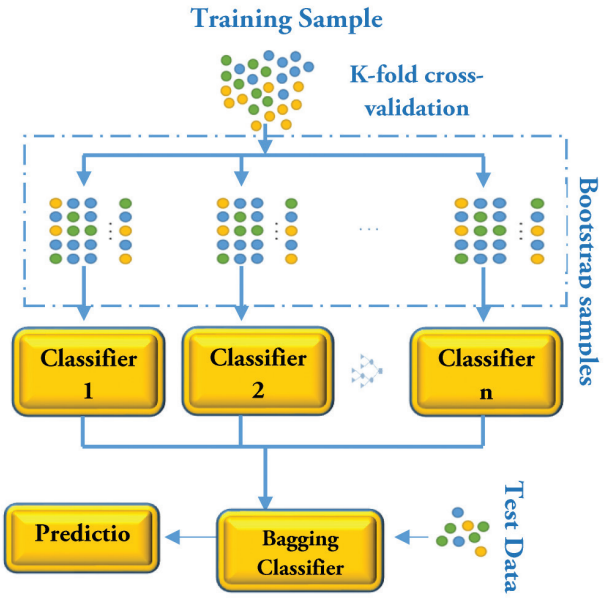


Fig. 5. Bagging process

In machine learning, Cross-validation is a blind technique used frequently to enhance model prediction and reduce bias. The BCI dataset has been divided randomly into k sets for each subject individually (k-fold cross-validation) [53]. In this study, 5-fold cross-validation has been used. One of these sets was utilized as a testing set, while the other four were used as training sets. This method is repeated five times, each time with a different set.

Previous work has suffered from weaknesses since it was designed as an independent algorithm. However, such defects can be overcome by combining different algorithms. Therefore, this work combines BT and RF algorithms to perform better than operating individually.

The BT algorithm reduces the pre-steps needed to obtain high-dimensional feature data, including dimension reduction and feature selections. Thus, from the speed evaluation perspective, the BT algorithm can speed up the process impressively while maintaining the easiness of creating parallel methods and the simplicity of implementation. Furthermore, if things go wrong, it can fix the error, as balancing errors occur by unbalanced data sets or trying to reduce the impact of losing a huge part of features by maintaining accuracy despite the current condition.

On the other hand, the RF algorithm focus on getting the best feature selection possible by calculating the Gini index and OOB error, which leads to judging the importance of features and interaction among different ones.

3.4.3. Smart Hybrid Algorithm

SBT–RF is implemented by combining the feature selection algorithm RF and classification algorithm BT to work with the BCI Data set. (Fig. 6) and Algorithm 4 shows the SBT-RF mechanism, starting with preparing each subject's dataset until it gets classified with 5-fold cross-validation.

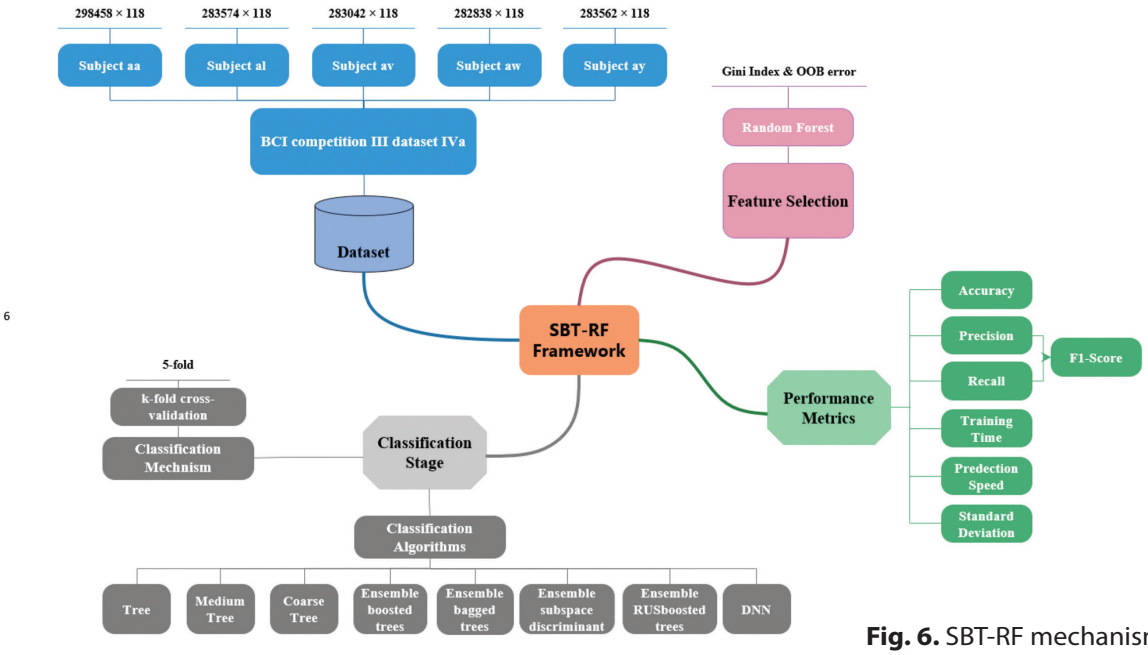


Fig. 6. SBT-RF mechanism

Algorithm 4: SBT-RF

Variables: OOB: Out of bag
OD: Original Dataset
N: Number of bootstrap samples
L: Learning Algorithm
C*: Bagging ensemble classifier
for each Subject
  for features x_j, j = 1 to z do
    for each tree learner in sample k = 1 to |k| do
      Find all nodes m that use x_j.
      Obtain feature importance by calculating impurity before and after splitting of x_j using the Gini index.
      Aggregate the enhancements
    end
    Aggregate the improvements of all the decision trees to the feature importance of x_j.
  end
  while developing tree
    Calculate accuracy for OOB observations and record the prediction
  end
  Do premutation to x_j OOB observations to breakdown the relations between the target & the permuted feature x_j.
  for each OOB observation
    calculate the accuracy and record the prediction again
  end
  Get the average of the performance reduction that had been happened due to permute x_j and utilize it to measure the x_j feature importance.
  Train BT algorithm using the selected feature with 5k-fold
  for j=1 to N
    BSSj ← bootstrap sample from OD
    Create Classifier Cj ←L(BSSj)
  end
  for each new instance, predict the class label
    $C^*(x) = arg \max_y \sum_{j=1}^{N} [C_j(x) = y]$
  end
  Evaluate the model performance
end

4. SIMULATION AND COMPUTER RESULTS

This section has three parts. The first section describes the dataset and how it was prepared. The second section discusses the performance metrics used. The final section displays the results of the proposed SBT-RF algorithm and other algorithms.

4.1. PERFORMANCE METRICS

Performance metrics indicate how the proposed method compares to each state-of-the-art algorithm, as shown in (Fig. 7).

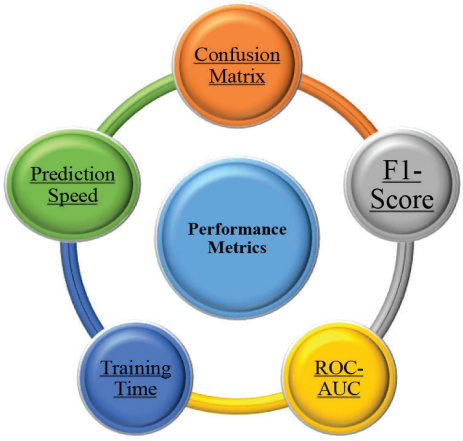


Fig. 7. Performance metrics.

Confusion Matrix: A Confusion Matrix is a table frequently used to specify the output of a classification model on a set of test data whose real values are known [54], [55]. It allows the performance of an algorithm to be visualized [56].

Accuracy and Misclassification rate equations are given as shown:

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (18)$$

$$\text{Misclassification rate} = \frac{FP+FN}{TP+TN+FP+FN} \quad (19)$$

Where TP: True Positive, TN: True Negative, FP: False Positive, FN: False Negative.

F1-Score: F1-Score is a technique to measure the performance depending on the harmonic way between Precision and Recall to grant a balanced measure of the misclassified cases than the confusion matrix [56].

$$\text{Precision} = \frac{\text{(True Positive)}}{\text{(True Positive + False Positive)}} \quad (20)$$

$$\text{Recall} = \frac{\text{(True Positive)}}{\text{(True Positive + False Negative)}} \quad (21)$$

From equations (20), (21); F1-Score can be calculated as shown:

$$F1 = \frac{2}{\text{(1/Precision + 1/Recall)}} \quad (22)$$

ROC-AUC: The Receiver Operating Characteristic Curve (ROC) is the plot that can show binary classifier performance as a function of the cut-off threshold between True positive rate (Sensitivity) and false positive rate (1-specificity) [57], [58]. Area Under Curve (AUC) supplies an overall performance measurement among every potential classification threshold [59].

Training Time: The overall time required for the model to learn.

Prediction Speed: The number of observations the machine learning model can produce per second.

4.2. COMPUTER SIMULATIONS AND RESULTS

The following section will describe the experiment that was conducted. First, putting several algorithms to use on BCI Competition III dataset IVa., then performing SBT-RF 50 times on each subject.

This section will present and apply several algorithms to BCI Competition III dataset IVa. All these algorithms were tested with 5-fold cross-validations to ensure the data's variance and reduce the computation cost. However, the algorithm does not need more K-folds because the BT algorithm uses bootstrap sampling to prevent the machine learning model from overfitting by sampling the dataset into random sets and training every group alone [18], [22].

Table 3 contains the average results of 50 runs on the BCI-competition III dataset IVa using eight different algorithms. The test was carried out using the paired device's resources. As the key performance indicators in our study, classification accuracy and F1-score in each run are calculated. Bolded values indicate the best values for each algorithm. The standard deviation (SD) was also calculated to evaluate the algorithm's stability.

Comparing several algorithms and evaluating performance metrics on the models for each subject on the dataset are shown in Table 3. Regarding training time, some algorithms perform better, such as the Coarse Tree with all subjects at 38.8s. Others overcome in predicting observation per second like medium in (aa and av subjects), Coarse Tree in (al, aw and ay subjects), and Fine Tree in (av and ay subjects). The best average prediction Speed value belongs to the Coarse Tree algorithm with 180000 obs/sec, but the accuracy and F1-score for these algorithms were extremely low. Therefore, they are not reliable for the machine learning model. Although the BT algorithm did not achieve a high value in prediction speed and training time, this algorithm dominates the others in accuracy and F1-score for all subjects with a mean accuracy of 97.64% farther than Deep Neural Network (DNN) by 15.31% and a mean F1-score 97.58% farther than DNN by 14.1%. While dealing with BCI datasets, the most important factor is reliability, so according to the results, the BT algorithm has been chosen to be a classification algorithm for this research. Now BT algorithm has good accuracy and F1 score. However, with average training time and prediction speed, a dimension reduction is made using feature selection to enhance computation cost, considering the same accuracy and F1-score or better values.

Table 3. Reported performance metrics.

Algorithm	Performance Metrics	Subjects					Mean
		aa	al	av	aw	ay	
Tree	Accuracy (%)	64.002	68.019	82.9	74.801	91.701	76.2846
	F1 score (%)	58.4	61.286	71.831	74.504	94.179	72.04
	Training Time (sec)	139	200	73	47	14	94.6
	Prediction Speed (obs/sec)	180000	180000	**170000**	140000	**180000**	170000
	Standard deviation (±%)	0.227	0.178	0.159	0.193	0.143	0.18
Medium Tree	Accuracy (%)	60.602	60.408	70.904	66.018	85.409	68.6682
	F1 score (%)	61.025	59.406	68.702	68.475	90.175	69.5566
	Training Time (sec)	105	120	44	27	10	61.2
	Prediction Speed (obs/sec)	**190000**	200000	**170000**	140000	170000	174000
	Standard deviation (±%)	0.157	0.108	0.112	0.15	0.127	0.1308
Coarse Tree	Accuracy (%)	57.702	56.706	67.304	60.009	76.802	63.7046
	F1 score (%)	59.921	42.751	61.206	66.775	84.794	63.0894
	Training Time (sec)	**65**	**73**	**31**	**18**	**7**	**38.8**
	Prediction Speed (obs/sec)	180000	**240000**	150000	**150000**	**180000**	**180000**
	Standard deviation (±%)	0.124	0.096	0.099	0.117	0.13	0.1132
Ensemble boosted trees	**Accuracy (%)**	**66.015**	**65.306**	**79.305**	**72.4**	**93.411**	**75.2874**
	F1 score (%)	60.233	64.53	62.165	73.789	95.401	71.2236
	Training Time (sec)	2604	3551	893	1385	194	1725.4
	Prediction Speed (obs/sec)	78000	70000	69000	140000	99000	91200
	Standard deviation (±%)	0.111	0.057	0.084	0.111	0.126	0.0978
Ensemble bagged trees	Accuracy (%)	**96.615**	**99.109**	**98.902**	**95.352**	**98.261**	**97.6478**
	F1 score (%)	**96.397**	**99.475**	**98.714**	**96.614**	**99.018**	**98.0436**
	Training Time (sec)	1652	2208	368	282	72	916.4
	Prediction Speed (obs/sec)	17000	21000	37000	27000	42000	28800
	Standard deviation (±%)	0.128	**0.06**	0.113	0.146	0.122	0.1138
Ensemble subspace Discriminant	Accuracy (%)	76.706	72.608	81.302	89.008	91.802	82.2852
	F1 score (%)	70.123	72.65	68.803	89.357	94.322	79.051
	Training Time (sec)	929	1207	163	202	47	509.6
	Prediction Speed (obs/sec)	4900	4300	7000	6500	7800	6100
	Standard deviation (±%)	0.208	0.167	0.173	0.162	0.223	0.1866

Algorithm	Performance Metrics	Subjects					Mean
		aa	al	av	aw	ay	
Ensemble RUSboosted trees	Accuracy (%)	60.709	60.576	72.905	70.808	91.206	71.2408
	F1 score (%)	65.163	59.512	67.086	68.598	93.783	70.8284
	Training Time (sec)	3131	4922	771	719	167	1942
	Prediction Speed (obs/sec)	89000	83000	66000	100000	110000	89600
	Standard deviation (±%)	0.126	0.112	0.118	0.116	0.146	0.1236
DNN	Accuracy (%)	76.24	92.36	64.57	87.68	90.83	82.336
	F1 score (%)	78.24	92.42	67.57	89.62	91.83	83.936
	Training Time (sec)	1199	990	115	244	212	552
	Prediction Speed (obs/sec)	115400	125200	139300	118500	145100	128700
	Standard deviation (±%)	**0.058**	**0.06**	0.086	**0.071**	**0.065**	**0.0682**

Boxplots in (Fig. 8) and (Fig. 9) reveal that the dataset's median is too near to its median after RF. The RF algorithm chose the right characteristics to describe the data well and rejected redundant ones [60], [61].

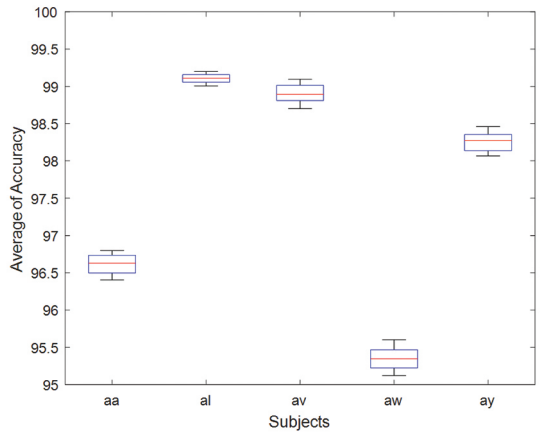


Fig. 8. 118-channel boxplot for BCI III Dataset IVa

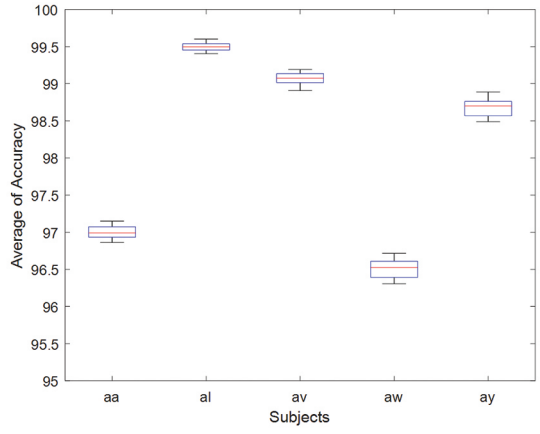


Fig. 9. RF algorithm selected channels boxplot for BCI III Dataset IVa

The second part of the experiment is executing SBT-RF on each subject 50 times. Again, two primary performance metrics (classification accuracy and F1-Score) were deployed to each run. (Fig. 10) and (Fig. 11) show that the SBT-RF algorithm is better than the BT algorithm's accuracy and the high F1 score, as shown in Table 4. Furthermore, (Fig. 12) delivers a decrease in training time (4582 to 1209) sec, while (Fig. 13) shows an increase in average prediction speed (28,800 to 70,600) obs/sec. Furthermore, the rate was increased by 12 times, the risk was limited (by reducing the electrodes) [62], and the computational and financial costs were decreased.

Table 4. SBT-RF vs. BT

Subject	Precision (%)	Recall (%)	F1-Score (%)	Accuracy (%)	Training Time (sec)	Prediction Speed (obs/sec)	AUC
BT							
aa	95.48	96.54	96.001	96.615	1652	17000	0.9
al	99.24	99.00	99.1	99.109	2208	21000	1
av	98.16	98.77	98.461	98.902	368	37000	1
aw	96.07	95.14	95.611	95.352	282	27000	0.9
ay	99.66	97.85	98.748	98.261	72	42000	1
SBT-RF							
aa	**95.91**	**97.04**	**96.397**	**97.003**	**475**	**46000**	**1**
al	**99.52**	**99.41**	**99.475**	**99.498**	**470**	**82000**	**1**
av	**98.48**	**98.96**	**98.714**	**99.073**	**133**	**92000**	**1**
aw	**97.07**	**96.18**	**96.614**	**96.512**	**98**	**59000**	**0.9**
ay	**99.69**	**98.39**	**99.018**	**98.682**	**33**	**74000**	**1**

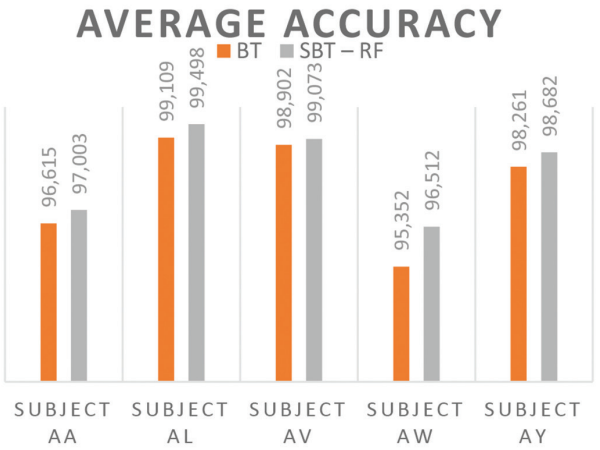


Fig. 10. Average accuracy for BT vs SBT-RF

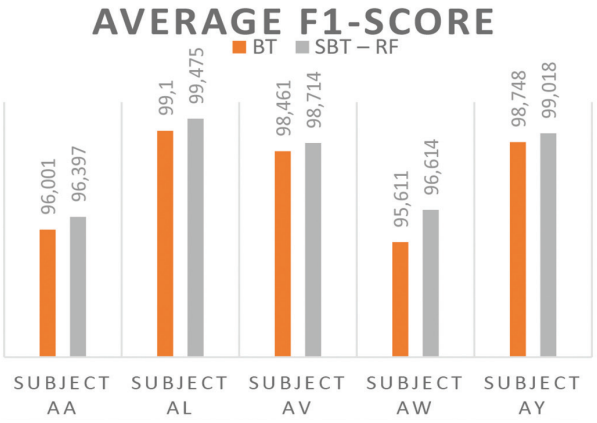


Fig. 11. Average F1-Score for BT vs SBT-RF

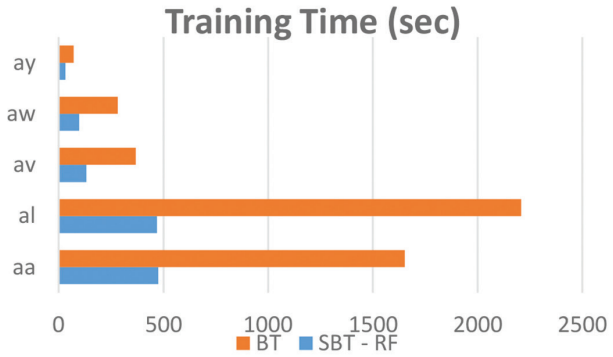


Fig. 12. Training time for BT vs SBT-RF

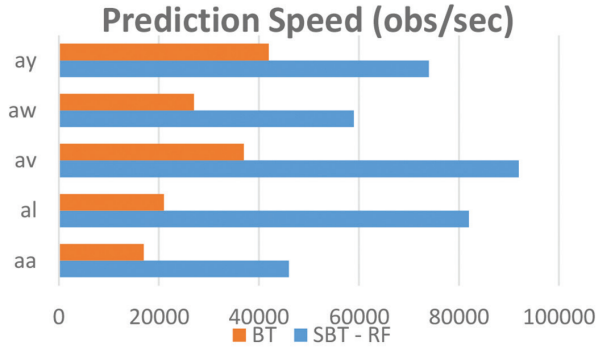


Fig. 13. Average prediction speed for BT vs SBT-RF

Table 5 compares the proposed method SBT-RF with related studies that use different algorithms on the same data set. SBT-RF outperforms all related work for av, aw, and ay subjects by 99.1%, 96.5%, and 98.7%, respectively. However, the enhanced CSP + LS-SVM method shows better results only in aa with 3%. While CSP+LDA and DSSMM methods achieve better results only in al, SBT-RF achieves the highest accuracy with 98.16% on average.

5. CONCLUSIONS

BCI is an advanced approach that helps analyze and recognize human intentions and thoughts that are further transformed into EEG signals. However, it is believed that some signals detected from certain brain channels may contain redundant data that decreases the classification efficiency. Accordingly, feature selection methods have been applied to evacuate the redundant data before the classification process to reduce computation costs. Furthermore, scientists usually combine different classifier combinations to overcome weak ones to make the classification algorithm more accurate. A smart hybrid algorithm (SBT-RF) for classifying BCI datasets (SBT-RF) is proposed in this work. The proposed algorithm is implemented in two stages; firstly, the RF algorithm evaluates the features' importance to select the most useful features. Then, this algorithm (i.e., RF) measures the extent of each feature individually by calculating the lowest impurity using the Gini index and obtaining OOB error. The selected features from the previous stage are inputs to the Ensemble BT classifier. Next, the BT classifier samples the processed "BCI Competition III Dataset IVa" into bootstrap samples, then classifies every sample individually using a decision tree. As a result, it prevents overfitting and collects the votes of the decision tree classifier to predict the class. It is revealed that the proposed algorithm has the highest average accuracy of ~98 % compared to other relevant algorithms reported in the literature.

Table. 5. SBT-RF vs related works

Author	Method	Subject Results					
		aa	al	av	aw	ay	Mean
Amin Hekmatmanesh et al. [28]	DFBCSP DSLVQ SSVM GRBF	93.5%	98.57%	81.78%	93.57%	96.07%	92.70%
Sahar Selim et al. [29]	CSP+LDA	79.46%	**100%**	57.14%	92.41%	69.84%	79.77%
Yao Guo et al. [4]	FCCSP	72.32%	98.21%	68.87%	78.57%	92.06%	82.01%
Wenlong Hang et al. [30]	DSSMM	75.89%	**100%**	76.53%	89.73%	76.19%	83.68%
Md.A.M. Joadder et al. [26]	Katz + LDA	86.78%	90.35%	68.92%	92.14%	83.57%	84.35%
Yongkoo Park et al. [27]	LRFCSP	98.93%	93.21%	81.79%	93.21%	97.5%	92.93%
Proposed Method	SBT-RF	97%	99.49%	**99.07%**	**96.51%**	**98.68%**	**98.15%**

6. REFERENCES:

[1] Y. You, W. Chen, T. Zhang, "Motor imagery EEG classification based on flexible analytic wavelet transform", Biomedical Signal Processing and Control, Vol. 62, 2020, p. 102069.

[2] J. Wang, Z. Feng, X. Ren, N. Lu, J. Luo, L. Sun, "Feature subset and time segment selection for the classification of EEG data based motor imagery", Biomedical Signal Processing and Control, Vol. 61, 2020, p. 102026.

[3] K. Sayed, M. Kamel, M. Alhaddad, H. M. Malibary, Y. M. Kadah, "Characterization of phase space trajectories for Brain-Computer Interface", Biomedical Signal Processing and Control, Vol. 38, 2017, pp. 55-66.

[4] Y. Guo, Y. Zhang, Z. Chen, Y. Liu, W. Chen, "EEG classification by filter band component regularized common spatial pattern for motor imagery", Biomedical Signal Processing and Control, Vol. 59, 2020, p. 101917.

[5] B. O. Olcay and B. Karaçalı, "Evaluation of synchronization measures for capturing the lagged synchronization between EEG channels: A cognitive task recognition approach", Computers in Biology and Medicine, Vol. 114, 2019, p. 103441.

[6] J. Jin, Y. Miao, I. Daly, C. Zuo, D. Hu, A. Cichocki, "Correlation-based channel selection and regularized feature optimization for MI-based BCI", Neural Networks, Vol. 118, 2019, pp. 262-270.

[7] P. Gaur, R. B. Pachori, H. Wang, G. Prasad, "A multiclass EEG-based BCI classification using multivariate empirical mode decomposition based filtering and Riemannian geometry", Expert Systems with Applications, Vol. 95, 2018, pp. 201-211.

[8] O. Landau, A. Cohen, S. Gordon, N. Nissim, "Mind your privacy: Privacy leakage through BCI applications using machine learning methods", Knowledge-Based Systems, Vol. 198, 2020, p. 105932.

[9] J. Feng et al. "Towards correlation-based time window selection method for motor imagery BCIs", Neural Networks, Vol. 102, 2018, pp. 87-95.

[10] A. Barachant, S. Bonnet, M. Congedo, C. Jutten, "Classification of covariance matrices using a Riemannian-based kernel for BCI applications", Neurocomputing, Vol. 112, 2013, pp. 172-178.

[11] Y. Li, P. P. Wen, "Clustering technique-based least square support vector machine for EEG signal classification", Computer methods and Programs in Biomedicine, Vol. 104, No. 3, 2011, pp. 358-372.

[12] N. Lu, T. Yin, "Motor imagery classification via combinatory decomposition of ERP and ERSP using sparse nonnegative matrix factorization", Journal of Neuroscience Methods, Vol. 249, 2015, pp. 41-49.

[13] J. Lee, J.-Y. Jeong, C.-H. Jun, "Markov blanket-based universal feature selection for classification and regression of mixed-type data", Expert Systems with Applications, Vol. 158, 2020, p. 113398.

[14] F. Haider, S. Pollak, P. Albert, S. Luz, "Emotion recognition in low-resource settings: An evaluation of automatic feature selection methods", Computer Speech & Language, Vol. 65, 2021, p. 101119.

[15] G. H. de Rosa, J. P. Papa, X.-S. Yang, "A nature-inspired feature selection approach based on hypercomplex information", Applied Soft Computing, Vol. 94, 2020, p. 106453.

[16] C. Li, X. Luo, Y. Qi, Z. Gao, X. Lin, "A new feature selection algorithm based on relevance, redundancy and complementarity", Computers in Biology and Medicine, Vol. 119, 2020, p. 103667.

[17] C.-F. Tsai, Y.-T. Sung, "Ensemble feature selection in high dimension, low sample size datasets: Parallel and serial combination approaches", Knowledge-Based Systems, Vol. 203, 2020, p. 106097.

[18] L. Breiman, "Random forests", Machine Learning, Vol. 45, No. 1, pp. 5-32, 2001.

[19] A. Zelaia, I. Alegria, O. Arregi, B. Sierra, "A multiclass/multilabel document categorization system: Combining multiple classifiers in a reduced dimension", Applied Soft Computing, Vol. 11, No. 8, 2011, pp. 4981-4990.

[20] P. Davis, F. Aziz, M. T. Newaz, W. Sher, L. Simon, "The classification of construction waste material using a deep convolutional neural network", Automation in Construction, Vol. 122, 2021, p. 103481.

[21] S. C. Lemon, J. Roy, M. A. Clark, P. D. Friedmann, W. Rakowski, "Classification and regression tree analysis in public health: methodological review and comparison with logistic regression", Annals of Behavioral Medicine, Vol. 26, No. 3, 2003, pp. 172-181.

[22] L. Breiman, "Bagging predictors", Machine Learning, Vol. 24, No. 2, 1996, pp. 123-140.

[23] R. E. Banfield, L. O. Hall, K. W. Bowyer, W. P. Kegelmeyer, "A comparison of decision tree ensemble creation techniques", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 29, No. 1, 2006, pp. 173-180.

[24] G. Martinez-Munoz, D. Hernández-Lobato, A. Suárez, "An analysis of ensemble pruning techniques based on ordered aggregation", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 31, No. 2, 2008, pp. 245-259.

[25] D. Zhang, X. Zhou, S. C. H. Leung, J. Zheng, "Vertical bagging decision trees model for credit scoring", Expert Systems with Applications, Vol. 37, No. 12, 2010, pp. 7838-7843.

[26] M. A. M. Joadder, S. Siuly, E. Kabir, H. Wang, Y. Zhang, "A new design of mental state classification for subject independent BCI systems", IRBM, Vol. 40, No. 5, 2019, pp. 297-305.

[27] Y. Park, W. Chung, "Frequency-optimized local region common spatial pattern approach for motor imagery classification", IEEE Transactions on Neural Systems and Rehabilitation Engineering, Vol. 27, No. 7, 2019, pp. 1378-1388.

[28] A. Hekmatmanesh, H. Wu, F. Jamaloo, M. Li, H. Handroos, "A combination of CSP-based method with soft margin SVM classifier and generalized RBF kernel for imagery-based brain computer interface applications", Multimedia Tools and Applications, Vol. 79, No. 25, 2020, pp. 17521-17549.

[29] S. Selim, M. Tantawi, H. Shedeed, A. Badr, "A Comparative Analysis of Different Feature Extraction Techniques for Motor Imagery Based BCI System", Proceedings of the Joint European-US Workshop on Applications of Invariance in Computer Vision, 2020, pp. 740-749.

[30] W. Hang, W. Feng, S. Liang, Q. Wang, X. Liu, K.-S. Choi, "Deep stacked support matrix machine based representation learning for motor imagery EEG classification", Computer Methods and Programs in Biomedicine, Vol. 193, 2020, p. 105466.

[31] L. He, Y. Hu, Y. Li, D. Li, "Channel selection by Rayleigh coefficient maximization based genetic algorithm for classifying single-trial motor imagery EEG", Neurocomputing, Vol. 121, 2013, pp. 423-433.

[32] D. Fattahi, B. Nasihatkon, R. Boostani, "A general framework to estimate spatial and spatio-spectral filters for EEG signal classification", Neurocomputing, Vol. 119, 2013, pp. 165-174.

[33] N. Lu, T. Li, J. Pan, X. Ren, Z. Feng, H. Miao, "Structure constrained semi-nonnegative matrix factorization for EEG-based motor imagery classification", Computers in Biology and Medicine, Vol. 60, 2015, pp. 32-39.

[34] Z. Qiu, J. Jin, H.-K. Lam, Y. Zhang, X. Wang, A. Cichocki, "Improved SFFS method for channel selection in motor imagery based BCI", Neurocomputing, Vol. 207, 2016, pp. 519-527.

[36] K. K. Kandaswamy, G. Pugalenthi, K.-U. Kalies, E. Hartmann, T. Martinetz, "EcmPred: Prediction of extracellular matrix proteins based on random forest with maximum relevance minimum redundancy feature selection", Journal of Theoretical Biology, Vol. 317, 2013, pp. 377-383.

[35] Y. Saeys, T. Abeel, Y. Van de Peer, "Robust feature selection using ensemble feature selection techniques", Proceedings of the Joint European Conference on Machine Learning and Knowledge Discovery in Databases, 2008, pp. 313-325.

[37] X. Li, W. Chen, Q. Zhang, L. Wu, "Building auto-encoder intrusion detection system based on random forest feature selection", Computers & Security, Vol. 95, 2020, p. 101851.

[38] E. Izquierdo-Verdiguier, R. Zurita-Milla, "An evaluation of Guided Regularized Random Forest for classification and regression tasks in remote sensing", International Journal of Applied Earth Observation and Geoinformation, Vol. 88, 2020, p. 102051.

[39] X. Ji, B. Yang, Q. Tang, "Seabed sediment classification using multibeam backscatter data based on the selecting optimal random forest model", Applied Acoustics, Vol. 167, 2020, p. 107387.

[40] X. Ma, X. Sun, "Sequence-based predictor of ATP-binding residues using random forest and mRMR-IFS feature selection", Journal of Theoretical Biology, Vol. 360, 2014, pp. 59-66.

[41] A. Srivastava, S. Ghosh, N. Anantharaman, V. K. Jayaraman, "Hybrid biogeography based simultaneous feature selection and MHC class I peptide binding prediction using support vector machines and random forests", Journal of Immunological Methods, Vol. 387, No. 1-2, 2013, pp. 284-292.

[42] L. Breiman, J. H. Friedman, R. A. Olshen, C. J. Stone, "Classification and regressiontrees, wadsworth statistics", Probability Series, Belmont, California: Wadsworth, 1984.

[43] C. Wang, A. Wang, J. Xu, Q. Wang, F. Zhou, "Outsourced privacy-preserving decision tree classification service over encrypted data", Journal of Information Security and Applications, Vol. 53, 2020, p. 102517.

[44] D. Niu, K. Wang, L. Sun, J. Wu, X. Xu, "Short-term photovoltaic power generation forecasting based on random forest feature selection and CEEMD: A case study", Applied Soft Computing, Vol. 93, 2020, p. 106389.

[45] M. Seera, C. P. Lim, "Online motor fault detection and diagnosis using a hybrid FMM-CART model", IEEE Transactions on Neural Networks and Learning Ssystems, Vol. 25, No. 4, 2013, pp. 806-812.

[46] J. Cho, X. Li, Z. Gu, P. U. Kurup, "Recognition of explosive precursors using nanowire sensor array and decision tree learning", IEEE Sensors Journal, Vol. 12, No. 7, 2011, pp. 2384-2391.

[47] A. Suárez, J. F. Lutsko, "Globally optimal fuzzy decision trees for classification and regression", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 21, No. 12, 1999, pp. 1297-1311.

[48] Y. Wu, Y. Ke, Z. Chen, S. Liang, H. Zhao, H. Hong, "Application of alternating decision tree with AdaBoost and bagging ensembles for landslide susceptibility mapping", Catena, Vol. 187, 2020, p. 104396.

[49] W. Chen, H. Hong, S. Li, H. Shahabi, Y. Wang, X. Wang, B. Bin Ahmad, "Flood susceptibility modelling using novel hybrid approach of reduced-error pruning trees with bagging and random subspace ensembles", Journal of Hydrology, Vol. 575, 2019, pp. 864-873.

[50] F. Li, C.-H. Chen, G. Xu, L. P. Khoo, Y. Liu, "Proactive mental fatigue detection of traffic control operators using bagged trees and gaze-bin analysis", Advanced Engineering Informatics, Vol. 42, 2019, p. 100987.

[51] H. Hong et al. "Landslide susceptibility mapping using J48 Decision Tree with AdaBoost, Bagging and Rotation Forest ensembles in the Guangchang area (China)", Catena, Vol. 163, 2018, pp. 399-413.

[52] Y. Himeur, A. Alsalemi, F. Bensaali, A. Amira, "Robust event-based non-intrusive appliance recognition using multi-scale wavelet packet tree and ensemble bagging tree", Applied Energy, Vol. 267, 2020, p. 114877.

[53] Z. Xiong, Y. Cui, Z. Liu, Y. Zhao, M. Hu, J. Hu, "Evaluating explorative prediction power of machine learning algorithms for materials discovery using k-fold forward cross-validation", Computational Materials Science, Vol. 171, 2020, p. 109203.

[54] A. Luque, A. Carrasco, A. Martín, A. de las Heras, "The impact of class imbalance in classification performance metrics based on the binary confusion matrix", Pattern Recognition, Vol. 91, 2019, pp. 216-231.

[55] D. Simon, D. L. Simon, "Analytic confusion matrix bounds for fault detection and isolation using a sum-of-squared-residuals approach", IEEE Transactions on Reliability, Vol. 59, No. 2, 2010, pp. 287-296.

[56] W. Castro, J. Oblitas, M. De-La-Torre, C. Cotrina, K. Bazán, H. Avila-George, "Classification of cape gooseberry fruit according to its level of ripeness using machine learning techniques and different color spaces", IEEE Access, Vol. 7, 2019, pp. 27389-27400.

[57] Y. Zhou et al. "Predictors of response to repeated ketamine infusions in depression with suicidal ideation: an ROC curve analysis", Journal of Affective Disorders, Vol. 264, 2020, pp. 263-271.

[58] A. Ben-David, "About the relationship between ROC curves and Cohen's kappa", Engineering Applications of Artificial Intelligence, Vol. 21, No. 6, 2008, pp. 874-882.

[59] G. Tripepi, K. J. Jager, F. W. Dekker, C. Zoccali, "Diagnostic methods 2: receiver operating characteristic (ROC) curves", Kidney International, Vol. 76, No. 3, 2009, pp. 252-256.

[60] F. Lei, P. Dupuis, O. Durrieu, G. Zissis, P. Maussion, "Acoustic resonance detection using statistical methods of voltage envelope characterization in metal halide lamps", IEEE Transactions on Industry Applications, Vol. 53, No. 6, 2017, pp. 5988-5996.

[61] Z. Li, N. Wang, Y. Li, X. Sun, M. Huo, H. Zhang, "Collective efficacy of support vector regression with smoothness priority in marine sensor data prediction", IEEE Access, Vol. 7, 2019, pp. 10308-10317.

[62] A. J. Fowle, C. D. Binnie, "Uses and abuses of the EEG in epilepsy", Epilepsia, Vol. 41, 2000, pp. S10-S18.

Automatic land cover classification with SAR imagery and Machine learning using Google Earth Engine

Original Scientific Paper

Geeta T. Desai
Department of Electronics and Telecommunication,
Babasaheb Naik College of Engineering, Pusad, Maharashtra, India.
tgeetadesai@gmail.com

Abhay N. Gaikwad
Department of Electronics and Telecommunication,
Babasaheb Naik College of Engineering, Pusad, Maharashtra, India.
abhay.n.gaikwad@gmail.com

Abstract *– Land cover is the most critical information required for land management and planning because human interference on land can be easily detected through it. However, mapping land cover utilizing optical remote sensing is not easy due to the acute shortage of cloud-free images. Google Earth Engine (GEE) is an efficient and effective tool for huge land cover analysis by providing access to large volumes of imagery available within a few days after acquisition in one consolidated system. This article demonstrates the use of Sentinel-1 datasets to create a land cover map of Pusad, Maharashtra using the GEE platform. Sentinel-1 provides Synthetic Aperture Radar (SAR) datasets that have a temporally dense and high spatial resolution, which is renowned for its cloud penetration characteristics and round-the-year observations irrespective of the weather. VV and VH polarization sentinel-1 time series data were automatically classified using a support vector machine (SVM) and Random Forest (RF) machine learning algorithms. Overall accuracies (OA), ranging from 82.3% to 90%, were obtained depending on polarization and methodology used. RF algorithm with VV polarization dataset stands better in comparison to SVM achieving OA of 90% and Kappa coefficient of 0.86. The highest user accuracy was obtained for the water class with both classifiers.*

Keywords*: Land cover classification, Google Earth Engine, Synthetic aperture radar, Random Forest and Support vector machine*

1. INTRODUCTION

Land cover (LC) is the physical surface of earth consisting of water, agriculture, soil, forest and other physical features of the earth's surface [1]. Availability of precise and timely global and regional level LC mapping information is vital for environmental monitoring, precision agriculture, urban planning and others [2-4]. LC classification and vegetation mapping information is crucial for policy making to help farmers for planning their agricultural resources. This makes the LC and vegetation mapping a key factor in environmental studies [5]. Due to the large scale and free availability of remotely sensed data, it has gained recognition as a prominent data source for LC mapping over the time [6]. In literature a number of LC classification studies are carried out using sentinel2 optical and sentinel1 radar sensors. LC mapping utilizing optical remote sensing is a difficult task due to acute shortage of cloud-free images [7]. Uninterrupted, round-the-clock observations of Sentinel-1A in all weather conditions make it a preferred data source for land monitoring, especially in areas having continuous cloud cover [8]. Sentinel-1A (S1A), operates with a revisit interval of 12-day, 20m spatial resolution, and with two polarizations vertical transmit, vertical receive (VV) and vertical transmit, horizontal receive (VH) [8].

Google Earth Engine (GEE) is recognized as a cloud-based computing platform having immense capacity for processing, storage and integration of satellite data. The availability of different state of art classifiers and fast processing speed have made GEE more popular among researchers [9].

Machine learning (ML) is an effective and efficient methodology for remote sensing applications. Over the last two decades there have been substantial advancements in developing ML based approaches for LC mapping using remote sensing imagery [10, 11]. Support vector machine (SVM) classifiers have received more attention in earth science applications due to their

efficient performance with limited training data [12]. Popularity of Random Forest (RF) has increased in the remote sensing field due to its excellent classification accuracy and processing speed [13]. The aim of the paper is to quantify the ability of Sentinel-1A dataset for the LC type mapping utilizing ML algorithms like SVM and RF on the Google Earth Engine (GEE) platform and compare their performance using user accuracy (UA), producer accuracy (PA) overall accuracy (OA) and F1 score. The rest of the article is structured as follows: In section 2 short review on LC mapping using SAR is covered.

In section 3 study area is introduced. In section 4 methodology used is presented while in section 5 experimental results, conclusion and future work are discussed.

2. RELATED WORK

Over the past decade, there is growing interest in the LC classification using remote sensing techniques due to its importance in different applications related to environmental studies. The SAR imagery has been used by many researchers for getting information about the type of land cover.

In [14] authors have presented the use of Sentinel-1 data to classify four classes of LC in the province of Punjab located in Pakistan, the results of the study concluded that joint use of coherence with backscattered intensity improves the accuracy of the classification to obtain 80% overall accuracy.

In [15] authors have highlighted that joint use of features from ascending and descending orbit dual-polarized images of sentinel-1 increases classification accuracy than using a single pass image.

In [16] authors have analysed the use of multisource and multitemporal SAR imagery to improve classification accuracy in wetlands using a multiple classifier system.

In [17] authors have used a combination of ascendant and descendant orbit images keeping both polarizations of SAR data for the finding of the built-up areas.

In [18] authors have evaluated the potential of sentinel-1A data by combining different attributes extracted from backscatter, polarimetry, and interferometry using RF and SVM classifiers for LC mapping of the Amazon region. Authors discriminated artificial surfaces, forests, and non-forested areas using multitemporal SAR data taken over small observation intervals of time. UA of 91% was obtained.

In [19] authors have highlighted the use of statistical properties of SAR images by using a statistical convolution neural network for LC classification.

In [20] authors have proposed the SAR LC classification approach which is based on region based classification to boost classification accuracy.

In [21] authors have confirmed the capabilities of interferometric SAR signatures for LC mapping in the mountainous area using k-principal component analysis and SVM.

In [22] authors have reported the use of Sentinel-1 interferometric coherence for LC mapping utilizing multiple feature based classifiers and concluded that the coherence information is important for LC mapping and gives better accuracies for all evaluated cases.

In [23] Patric H. et.al introduced Eurosat novel dataset for LC classification consisting of sentinel-2 images. The dataset covers 13 spectral bands made up of total 27,000 labelled and geo-referenced images for 10 classes.

Generating large-scale land cover maps requires multitemporal satellite data which gives rise to a tremendous amount of data [24]. Processing of big data needs massive storage capacity and access to powerful computational capability. This process becomes a tedious and time consuming task if traditional methods for image collection, filtering, downloading, and preprocessing are used [25]. To support this requirement huge amount of storage, access to high power computing and flexibility to implement diverse applications are required. All the above requirements were taken care of after the emergence of the open access GEE platform.

GEE, is an integrated cloud-based computing platform that comes with a powerful capability of Google and can resolve the prominent issues related to the LC mapping of huge study regions [9]. Users can process large volumes of remote sensing data without the need to download it to their machine on GEE web-based Integrated Development Environment code editor [25]. The Availability of different algorithm packages and fast processing makes GEE more popular among researchers [26].

GEE has become an asset to the remote sensing community by making satellite data processing very convenient and fast.

Researchers are utilizing GEE in recent years for LC classification. The use of GEE for land cover classification using Landsat8 [24,25,27,28], sentinel2 [29] and combinations of sentinel2 and landsat8[30] has shown good results. Therefore GEE presents great opportunities in dealing with remote sensing data for LC mapping in Pusad. Till date, most studies using GEE for LC mapping are focused on using Landsat and sentinel2 optical images. There are few studies that used SAR imagery for large-scale LC mapping using ML methods on the GEE platform.

3. STUDY AREA

Pusad is a small city in the central area of Maharashtra state in India, which lies on latitude 19.912676, and longitude 77.566910 and includes the Pus River.

Pusad has a total area of approximately 1176 sq. kms. The prime LC in the study area is agriculture and the prominent crops grown are wheat, soybean, cotton, and turmeric, extending over a total 65% of the area under investigation. In addition to agriculture, the study area also contains urban areas, trees and water bodies. All agricultural areas are grouped into a single class called agriculture. Other remaining classes are urban area, vegetation, water and bare soil. Fig. 2 shows a backscatter image of the Pusad region in VV and VH polarization.

4. METHODOLOGY

Fig.3 flowchart describes the methodology used for exploring the LC classification in the study area. The methodology is organized into three stages. In the first stage, SAR images of the study area are acquired and pre-processed. The second stage deals with defining classification classes and the selection of training samples. In the last stage, classification and evaluation of the algorithm are performed.

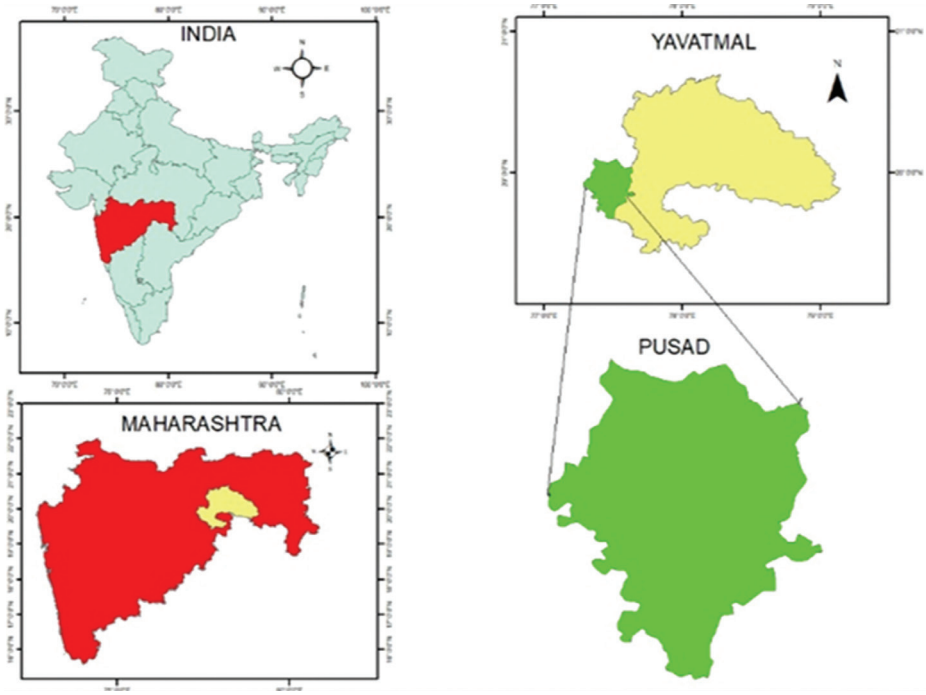


Fig. 1. Location Map of Pusad, India

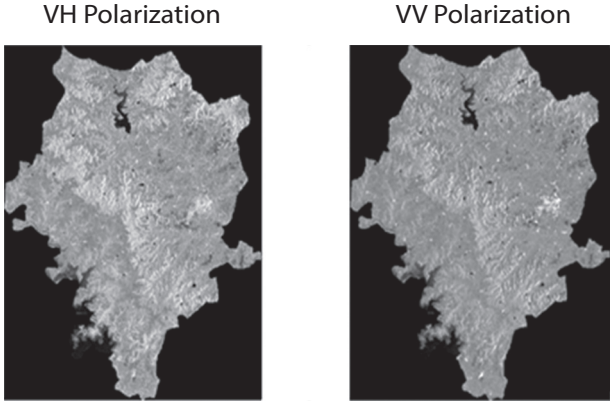


Fig. 2. Backscatter image of study area for VV and VH band

4.1 SENTINEL1 SAR IMAGE AND PREPROCESSING

In this study, the sentinel1- SAR GRD dataset of the study area stored in the GEE cloud platform was used and it included all images covering the area of investigation in the time period from January to April of the years 2018, 2019 and 2020. The Sentinel-1 SAR GRD dataset was acquired using the interference wide-band (IW) mapping mode, with a spatial resolution of 20 m, a width of 250 km, and an average incidence angle of 30–45 with a 12 day revisit time. The dataset contains VV and VH polarization. Sentinel-1 image preprocessing was implemented using GEE. Preprocessing includes orbit restitution, thermal noise removal, terrain correction and radiometric calibration. Each sentinel1 image was filtered to reduce speckle on the GEE platform [31].

4.2 TRAINING CLASSIFIER

Five types of LC i.e. agriculture, barren land, urban, vegetation, and water are the dominant part of the study area. Training and validation data was selected depending on human visual interpretation of high-resolution images from GEE [25,32]. Two popular supervised classifiers SVM and RF were deployed for LC mapping in Pusad.

4.2.1 SVM

SVM is a type of supervised learning, non-parametric classifier algorithm mostly used due to its ability to

work with limited training datasets with good classification results [33,34]. SVM training classifiers find an ideal hyperplane in the training phase that divides the dataset into predefined numerous classes with a few misclassified pixels. SVM uses support vectors to create the hyperplane. Support vectors are selected based on the cost parameter, Gamma, and kernel functions. The grid search based optimization technique is used to select C and Gamma parameters, producing trustworthy prediction outcomes. The random basis kernel is preferred for training on large datasets [33].

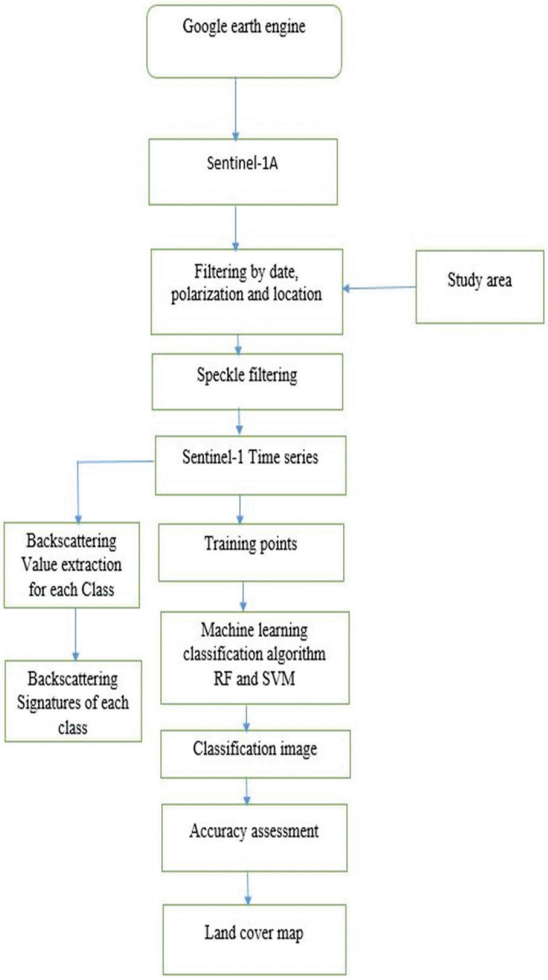


Fig. 3. Flowchart of methodology

4.2.2. Random Forest

RF is a type of ensemble classification technique. RF is found appealing to researchers in remote sensing over the last decade due to the identification and better handling of outliers in training samples using sample proximity measurement techniques [35,36].RF performs well with highly correlated hyperspectral data as compared to other streamlined ML classifiers. Another factor that makes RF a favorite of researchers is that it requires the setting of only two parameters ntree and mtry[10,37]. Taking into account all the above reasons we chose RF as one of the classification algorithms. As per suggestions of previous studies, we have set ntree parameter to 100 and mtry was set to the default value.

After completion of classification, an assessment step was performed for evaluating classification algorithms. Out of the total dataset, 70 percent was used for training the algorithm and the remaining 30 percent was used as a testing dataset. The OA, Kappa coefficient, UA and PA of each class was calculated using the inbuilt algorithm from GEE.

5. RESULTS AND DISCUSSIONS

5.1 TEMPORAL BACKSCATTER PATTERNS OF DIFFERENT LAND COVERS

Fig. 4 summarizes temporal backscatter signature curves of VV and VH bands for each land cover type in 2020. In the graph, the date on which the image is acquired is plotted on X-axis and backscattering coefficient is plotted on Y-axis. The average backscatter value was extracted from each image belonging to each LC class. From Fig. 4 it is observed that

i) Lowest backscatter values in VV and VH bands were noticed in the case of water bodies. Moreover, the urban class showed the maximum backscatter values for VV and VH bands.

(ii)Vegetation showed constant backscatter values throughout the time of image acquisition with slight variation in the month of September and October. The backscatter of the agricultural class showed more variation with VH.

(iii) Backscattering curves of vegetation and barren land showed similar shapes therefore identification of these classes was difficult.

5.2 LAND COVER CLASSIFICATION OF PUSAD

This study examines the potential of SVM and RF classifiers for LC classification using SAR images on the GEE platform. Fig. 5 and 6 present the classification results of VV and VH polarization using SVM and RF techniques for the year 2018. The accuracy assessment parameters showed variation in results with polarization and classifier.

In several studies [15 -23] authors processed remote sensing data on third party software, we implemented the entire methodology inside the GEE platform and used the JavaScript language for programming in the code editor. Analysing high volume data with remote sensing software on a personal computer is a tedious task however it is easily manageable on GEE. Hence GEE offers an attractive solution to users having low end devices to obtain satisfactory results in less time without the need of high power computers or commercial software. Most researchers previously used Landsat [24,25,27,28], Sentinel-2 [29] and a combination of Landsat and sentinel-2[30] images for generating LC classification. We chose sentinel-1 images since SAR sensors can acquire an image in all weather conditions in addition to these different combinations of wavelength and polarization provides important information regarding the earth's surface [26].

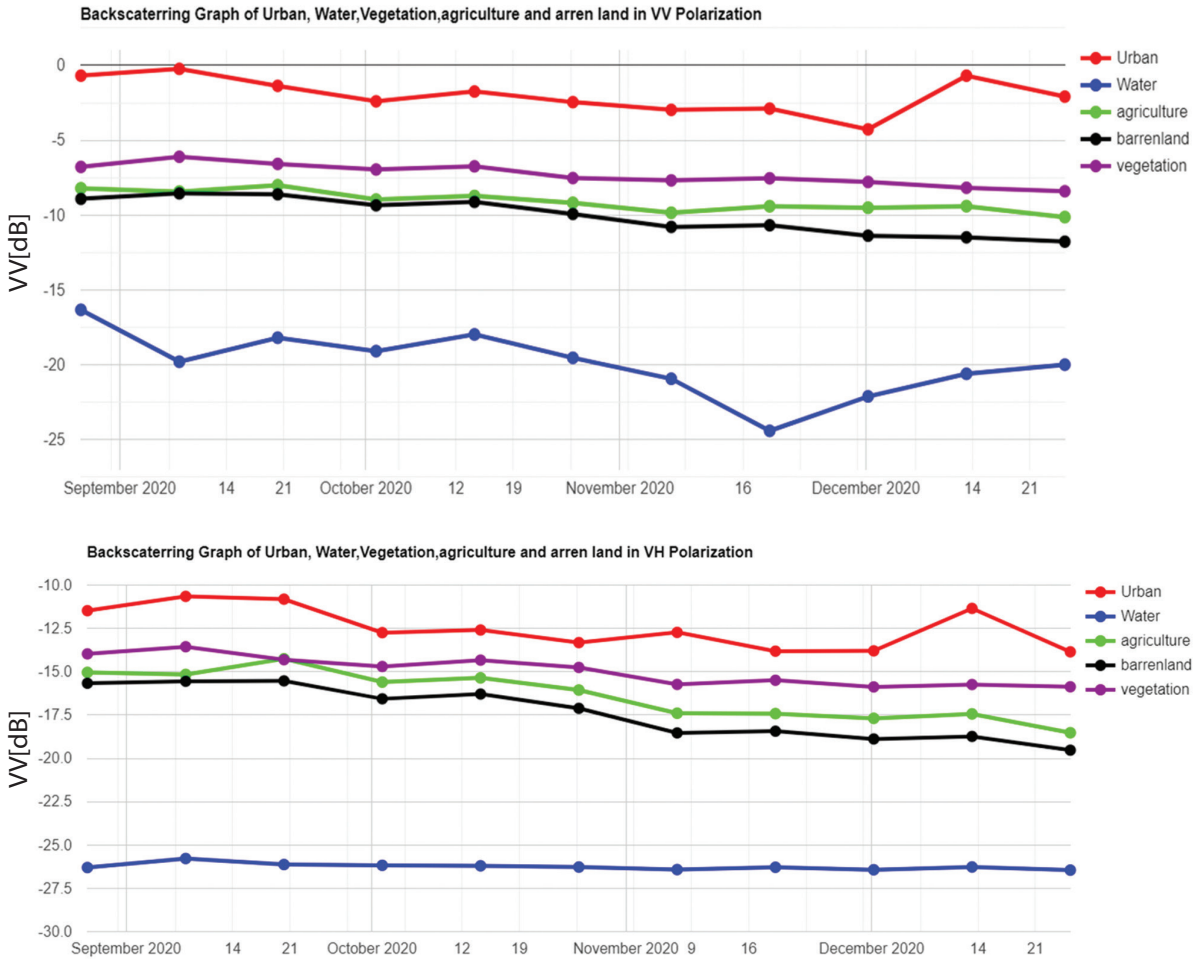
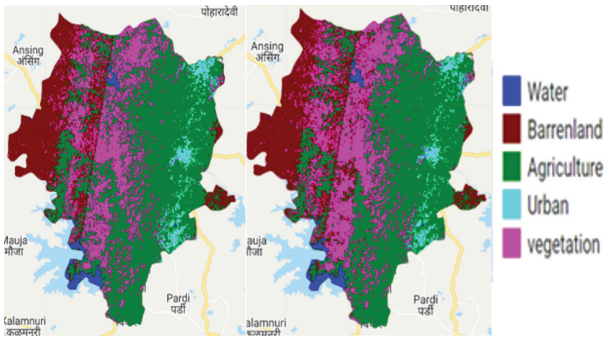
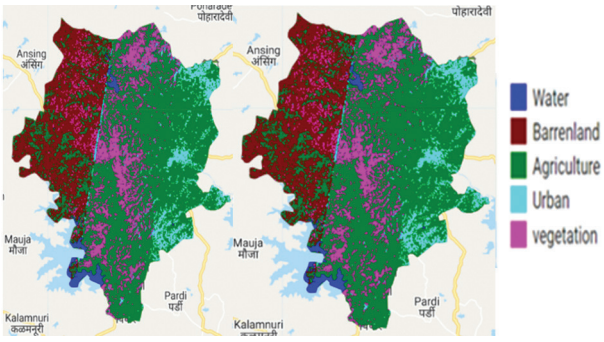


Fig. 4. Temporal backscatter profiles for five different land cover classes for VV and VH



(a) (b)

Fig. 5 . LC output Using RF for year 2018 for VH(a) and VV(b) band



(a) (b)

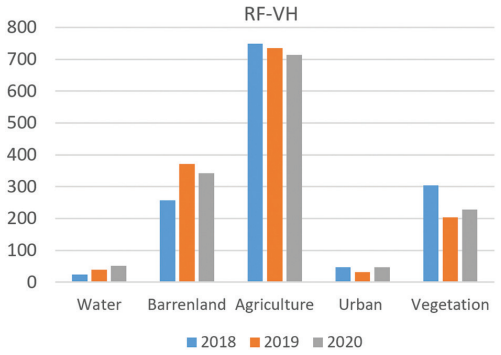
Fig. 6. LC output using SVM for year 2018 for VH(a) and VV(b) band

Table1, Table 2, Table 3 and Table 4 show the results of the SVM and RF classifier. UA, and PA, F1 score were used to access the class wise performance of each classifier. Both the classifiers get confused between barren land and vegetation areas however their performance for water, agriculture, and urban class is better. From the classification result it can be observed that most of the barren land is getting misclassified as an urban area.

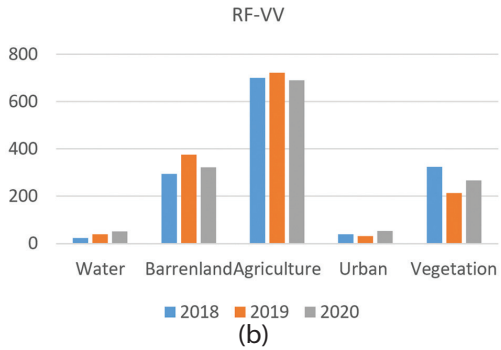
Our results show that RF has proved superior to the SVM classifier in line with results obtained in [26] Taking into account accuracies summarized in Table 2, it is clear that barren land and vegetation area are mostly misclassified classes. For SVM and RF, the performance of the water class was best as compared to other classes in agreement with the results obtained in [30]. Obtained results claim that VV polarization performs better in comparison to the VH band and similar results were found in [16]. From results it is observed that RF outperforms SVM in LC applications due to its robustness to outliers and noise also as compared to RF SVM is more sensitive

to hyperparameters [38]. Obtained results show that the presented approach achieves improvement of 3% in overall accuracy for land cover classification compared to obtained in [28] since SAR sensors can acquire an image in all weather conditions in addition to these different combinations of polarization provides important LC details of the earth's surface which improves LC classification accuracy.

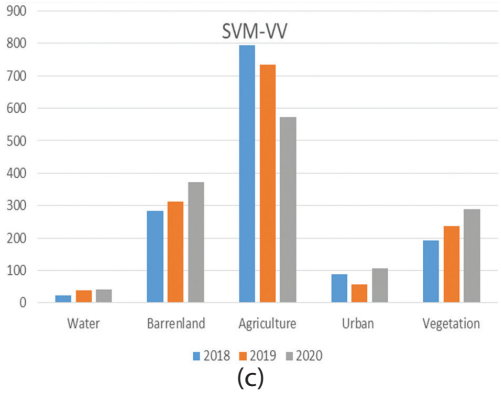
As shown in the Fig. 7, SVM achieved an accuracy of OA between 86.49% and 82% and kappa between 0.75 and 0.81 for the images from 2018 to 2020. While the RF achieved accuracy of OA between 90% and 85.8%. and kappa between 0.80 and 0.86 for the images from 2018 to 2020. The average kappa coefficient for SVM and RF was 0.78 and 0.83 respectively. The dominant LC in the study area is agriculture, barren land and vegetation. Considering the results of RF algorithm for VV polarization it can be concluded that from 2018 to 2020, urban coverage and barren land showed a commutative increase by 1.1% and 2.3% of the study area respectively. From 2018 to 2020 vegetation area including forests and grasslands showed a reduction of 4.9% of the study area. Agricultural was seen most dominant class in 2019 as compared to 2020.



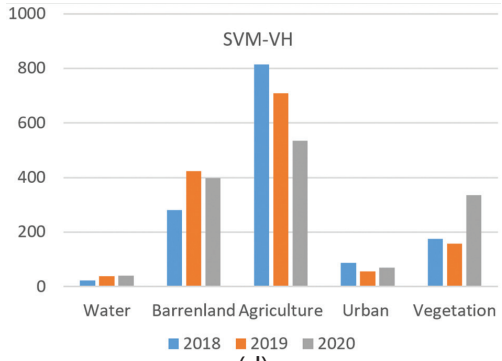
(a)



(b)



(c)



(d)

Fig. 7. Areas of water, Barrenland, Agriculture,urban and vegetation for 2018, 2019 and 2020.

Table 1. Comparison between SVM and RF classifier using Kappa coefficient and overall accuracy.

Year	Classifier	VV Band		VH Band	
		OA	Kappa Coefficient	OA	Kappa Coefficient
2018	SVM	86.49	0.818	85.3	0.79
	RF	90	0.86	89	0.85
2019	SVM	85.43	0.80	82.3	0.75
	RF	89	0.84	88	0.84
2020	SVM	85.71	0.805	83.27	0.77
	RF	85.8	0.80	87.5	0.83

Table 2. User accuracy and producer accuracy of each class for 2018 using RF and SVM

Class	RF		SVM	
	PA	UA	PA	UA
Water	93	97	98.60	94
Barren land	81	75	62.35	54
Agriculture	94.59	92	89.64	94.68
Urban	94.55	97	91.56	93.82
Vegetation	81.25	81	79.50	75.78

Table 3. User accuracy and producer accuracy of each class for 2018 using RF and SVM

Class	VV Band			VH Band		
	Precision	Recall	F1	Precision	Recall	F1
Water	97	93	95	94	99	96
Barren land	75	81	78	71	80	75
Agriculture	92	94.59	93	94	93	93
Urban	97	94.44	96	100	94	96
Vegetation	81	81.25	81	83	77	79

Table 4. Precision, Recall andF1 score of each class for 2018 using SVM classifier

Class	VV Band			VH Band		
	Precision	Recall	F1	Precision	Recall	F1
Water	97	93	95	94	99	96
Barren land	75	81	78	71	80	75
Agriculture	92	94.59	93	94	93	93
Urban	97	94.44	96	100	94	96
Vegetation	81	81.25	81	83	77	79

6. CONCLUSION:

The study evaluated the performance of RF and SVM for LC classification using time series Sentinel-1A datasets on the GEE platform for Pusad region. LC classification using SVM with VV gave overall accuracy of 85.87% with a kappa coefficient of 0.80, while classification with VH band gave an accuracy of 83.62% with a kappa coefficient of 0.77. With RF classifier VV gave overall accuracy of 88.26% with a kappa coefficient of 0.84 while with VH band overall accuracy of 88.1% with Kappa coefficient of 0.84 is obtained Based on the results it can be concluded that RF classifier using VV band gave best results. Agricultural and barren land is the main land cover types in the Pusad region. Future work can be focused on deep learning methodologies for LC mapping.

7. REFERENCES:

[1] J. S. Rawat, M. Kumar, "Monitoring land use/cover change using remote sensing and GIS techniques: A case study of Hawalbagh block, district Almora, Uttarakhand, India", The Egyptian Journal of Remote Sensing and Space Science, Vol. 18, No. 1, 2015, pp. 77-84.

[2] C. Zhang, J. M. Kovacs, "The application of small unmanned aerial systems for precision agriculture: a review", Precision agriculture, Vol. 13, No. 6, 2012, pp. 693-712.

[3] H. Shi et al. "Accurate urban area detection in remote sensing images", IEEE Geoscience and Remote Sensing Letters, Vol. 12, No. 9, 2015, pp. 1948-1952.

[4] Y. A. Liou, A. K. Nguyen, M. H. Li, "Assessing spatiotemporal eco-environmental vulnerability by Landsat data", Ecological Indicators, Vol. 80, 2017, pp. 52-65.

[5] R. S. DeFries, J. R. G. Townshend, "NDVI-derived land cover classifications at a global scale", International journal of remote sensing, Vol. 15, No. 17, 1994, pp. 3567-3586.

[6] N. Sidhu, E. Pebesma, G. Câmara, "Using Google Earth Engine to detect land cover change: Singapore as a use case", European Journal of Remote Sensing, Vol. 51, No. 1, 2018, pp. 486-500.

[7] C. Liu, Z.-X. Chen, S. H. A. O. Yun, J. Chen, T. Hasi, H. Pan, "Research advances of SAR remote sensing for agriculture applications: A review", Journal of integrative agriculture, Vol. 18, No. 3, 2019, pp. 506-525.

[8] P. Potin et al. "M.Sentinel-1 mission status", Proceedings of the EUSAR 2016: 11th European Conference on Synthetic Aperture Radar, Hamburg, Germany, 6-9 June 2016, pp. 1–6.

[9] N. Gorelick et al. "Google Earth Engine: Planetary-scale geospatial analysis for everyone", Remote Sensing of Environment, Vol. 202, 2017, pp. 18-27.

[10] A. E. Maxwell, T. A. Warner, F. Fang, "Implementation of machine-learning classification in remote sensing: An applied review", International Journal of Remote Sensing, Vol. 39, No. 9, 2018, pp. 2784-2817.

[11] D. J. Lary, A. H. Alavi, A. H. Gandomi, A. L. Walker, "Machine learning in geosciences and remote sensing", Geoscience Frontiers, Vol. 7, No. 1, 2016, pp. 3-10.

[12] P. Mantero, G. Moser, S. B. Serpico,"Partially supervised classification of remote sensing images through SVM-based probability density estimation", IEEE Transactions on Geoscience and Remote Sensing, Vol. 43, No. 3, 2005, pp. 559-570.

[13] V. F. Rodriguez-Galiano, B. Ghimire, J. Rogan, M. Chica-Olmo, J. P. Rigol-Sanchez, "An assessment of the effectiveness of a random forest classifier for land-cover classification", ISPRS journal of photogrammetry and remote sensing, Vol. 67, 2012, pp. 93-104.

[14] R. Z. Khalil, "InSAR coherence-based land cover classification of Okara, Pakistan", The Egyptian Journal of Remote Sensing and Space Science, Vol. 21, 2018, pp S23-S28.

[15] S. A. Sayedain, Y. Maghsoudi, S. Eini-Zinab, "Assessing the use of cross-orbit Sentinel-1 images in land cover classification", International Journal of Remote Sensing, Vol. 41, No. 20, 2020, pp. 7801-7819.

[16] M. Amani, B. Salehi, S. Mahdavi, B. Brisco, M. Shehata, "A Multiple Classifier System to improve mapping complex land covers: A case study of wetland classification using SAR data in Newfoundland, Canada", International Journal of Remote Sensing, Vol. 39, No. 21, 2018, pp. 7370-7383.

[17] I.-H. Holobâcă, K. Ivan, M. Alexe, "Extracting built-up areas from Sentinel-1 imagery using land-cover classification and texture analysis", International Journal of Remote Sensing, Vol. 40, No. 20, 2019, pp. 8054-8069.

[18] J. M. F. Diniz, F. F. Gama, M. Adami, "Evaluation of polarimetry and interferometry of sentinel-1A SAR data for land use and land cover of the Brazilian Amazon Region", Geocarto International, Vol. 37, No. 5, 2022, pp. 1482-1500.

[19] X. Liu, C. He, Q. Zhang, M. Liao, "Statistical convolutional neural network for land-cover classification from SAR images", IEEE Geoscience and Remote Sensing Letters, Vol. 17, No. 9, 2019, pp. 1548-1552.

[20] M. D. Soares et al. "A meta-methodology for improving land cover and land use classification with SAR imagery", Remote Sensing, Vol. 12, No. 6, 2020, pp. 961.

[21] H.W. Yun, J.-R. Kim, Y.-S. Choi, S.-Y. Lin, "Analyses of time series InSAR signatures for land cover classification: case studies over dense forestry areas with L-band SAR images", Sensors, Vol. 19, No. 12, 2019, pp. 2830.

[22] A. Jacob et al. "Sentinel-1 InSAR coherence for land cover mapping: A comparison of multiple feature-based classifiers", IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, Vol. 13, 2020, pp. 535-552.

[23] P. Helber, B. Bischke, A. Dengel, D. Borth, "Eurosat: A novel dataset and deep learning benchmark for land use and land cover classification", IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, Vol. 12, No. 7, 2019, pp. 2217-2226.

[24] H. Shafizadeh-Moghadam, M. Khazaei, S. K. Alavipanah, Q. Weng, "Google Earth Engine for large-scale land use and land cover mapping: an object-based classification approach using spectral, textural and topographical factors", GIScience & Remote Sensing, Vol. 58, No. 6, 2021, pp. 914-928.

[25] T. N. Phan, V. Kuch, L. W. Lehnert, "Land Cover Classification using Google Earth Engine and Random Forest Classifier—The Role of Image Composition", Remote Sensing, Vol. 12, No. 15, 2020, pp. 2411.

[26] H. Tamiminia, B. Salehi, M. Mahdianpari, L. Quackenbush, S. Adeli, B. Brisco,"Google Earth Engine for geo-big data applications: A meta-analysis and systematic review", ISPRS Journal of Photogrammetry and Remote Sensing, Vol. 164, 2020, pp. 152-170.

[27] N. Wagle et al. "Multi-temporal land cover change mapping using google earth engine and ensemble learning methods", Applied Sciences, Vol. 10, No. 22, 2020, p. 8083.

[28] Md Masroor et al. "Assessing the Influence of Land Use/Land Cover Alteration on Climate Variability: An Analysis in the Aurangabad District of Maharashtra State, India", Sustainability, Vol. 14, No.2, 2022, p. 642.

[29] S. Praticò, F. Solano, S. Di Fazio, G. Modica, "Machine learning classification of mediterranean forest habitats in google earth engine based on seasonal sentinel-2 time-series and input image composition optimisation", Remote Sensing, Vol. 13, No.4, 2021, pp. 586.

[30] K. N. Loukika, V. R. Keesara, V. Sridhar, "Analysis of land use and land cover using machine learning algorithms on google earth engine for Munneru River Basin, India", Sustainability, Vol. 13, No.24, 2021, pp. 13758.

[31] A. Mullissa et al. "Sentinel-1 sar backscatter analysis ready data preparation in google earth engine", Remote Sensing, Vol. 13, No.10, 2021, pp. 1954.

[32] M. C. Hansen et al. "A method for integrating MODIS and Landsat data for systematic monitoring of forest cover and change in the Congo Basin", Remote Sensing of Environment, Vol. 112, No.5, 2008, pp.2495-2513.

[33] G. Mountrakis, J. Im, C. Ogole, "Support vector machines in remote sensing: A review", ISPRS Journal of Photogrammetry and Remote Sensing, Vol. 66, No.3, 2011, pp.247-259.

[34] B.-E. Jamsran et al. "Applying a support vector model to assess land cover changes in the Uvs Lake Basin ecoregion in Mongolia", Information Processing in Agriculture Vol. 6, No.1, 2019, pp.158-169.

[35] X. Li et al., "A comparison of machine learning algorithms for mapping of complex surface-mined and agricultural landscapes using ZiYuan-3 stereo satellite imagery", Remote Sensing, Vol. 8, No.6, 2016, pp.514.

[36] Y. Jin et al. "Land-cover mapping using Random Forest classification and incorporating NDVI time-series and texture: A case study of central Shandong", International Journal of Remote Sensing, Vol. 39, No.23, 2018, pp.8703-8723.

[37] M. Belgiu, L. Drăguţ, "Random forest in remote sensing: A review of applications and future directions", ISPRS Journal of Photogrammetry and Remote Sensing, Vol. 114, 2016, pp.24-31

Selection of PID controller design plane for time-delay systems using genetic algorithm

Original Scientific Paper

Aye Taiwo Ajiboye
University of Ilorin,
Faculty of Engineering and Technology,
Department of Computer Engineering,
Ilorin, Kwara State, Nigeria.
ajiboye.at@unilorin.edu.ng

Olusogo Joshua Popoola
University of Ilorin,
Faculty of Engineering and Technology,
Department of Computer Engineering,
Ilorin, Kwara State, Nigeria.
olusogo@unilorin.edu.ng

Jayeola Femi Opadiji
University of Ilorin,
Faculty of Engineering and Technology,
Department of Computer Engineering,
Ilorin, Kwara State, Nigeria.
jopadiji@unilorin.edu.ng

Olalekan Femi Adebayo
University of Ilorin,
Faculty of Engineering and Technology,
Department of Computer Engineering,
Ilorin, Kwara State, Nigeria.
adebayo.of@unilorin.edu.ng

Abstract *– The design of a Proportional-Integral-Derivative (PID) controller with proportional, integral, and derivative, gain, k_p, k_i, and k_d, respectively, for a time-delay system, is quite common, particularly in the k_i- k_d plane, for a fixed k_p or in the k_p - k_i plane, for a fixed k_d. These design methods have been widely reported in the literature, however, the process of investigating the effects of using any of these design planes on system performance has not been given serious attention hence the need for this study. The stability region in the k_i- k_d and k_p - k_i design plane for a fixed value of k_p and k_d respectively were determined. For every determined stability region, the optimum value of controller gains in the plane was determined using a genetic algorithm (GA) with the integral of time multiplied by absolute error (ITAE) used as the objective function. The optimum value of the fixed gains was graphically determined by plotting the minimum of ITAE (Min-ITAE) for each stability region against the fixed gains. The overall optimum controller gains are the fixed gain that gives minimum of Min-ITAE (Min (Min-ITAE)) and the gains that resulted in Min-ITAE that yielded the Min (Min-ITAE). Using the determined overall optimum controller gains, the system closed-loop step response was plotted for the two design planes and the time domain performance measures (TDPMs) were determined. Based on TDPMs obtained for examples 1, 2, and 3, the k_i-k_d design plane yielded a faster response while the k_p- k_i design plane yielded a response that closely tracks the input irrespective of the system type and order. The study will enable control system designers to select the design plane that will give the best system performance right from the start of controller design without involving trial and error once the system transfer function and design specifications are known.*

Keywords*: Genetic algorithm, ITAE, PID controller, Stability region, Time-delay system, Time domain performance measures*

1. INTRODUCTION

All practical control systems have associated time delays [1-3] resulting from the processing and transmission of signals in the control loop [4, 5]. Therefore, the designing of a controller for such a system requires a full understanding of the effects of this delay on the system's performance. Time delay can degrade the system's quality of performance or destabilise the system in the worst case [6]. To control this class of systems, a PID controller is normally employed due to its popularity [7], simplicity, robustness, and ease of use [8-10]. To design a PID controller for a time-delay system the first step is to establish the stability boundary in the space of the controller parameter [11]. This is because of the lots of work that have been done in this area to the most recent, designs of PID controller for time-delay systems are normally carried out using the stability locus method [12].

One of the proven steps for the design of a PID controller for time-delay systems using the stability locus method involves plotting the stability boundaries in the k_i-k_d plane for a fixed k_p or in the k_p-k_i plane for a fixed k_d and then computing all the stabilizing values of k_p, k_i, and k_d in the stability region by sweeping over k_p and k_d for k_i-k_d, and k_p-k_i design planes respectively [16, 17].

The method for computation of all the stabilizing PID controller gains for a linear arbitrary order system with time delay in the k_i-k_d plane with fixed k_p was re-

ported in [7], while the procedure for computing the entire stability gains in the k_p-k_i plane with fixed k_d was explained in [13]. The stabilization of controller parameters in the k_p-k_i plane for a given k_d, for a time-delay integral fractional-order system, under the control of a fractional-order PID controller, was earlier determined and reported in [14]. It was observed that the controller gains in the k_p-k_i plane for a given k_d and other controller parameters yielded a general stability region, while the global stability region was obtained after the griding of k_d. PID parameters for a second-order plant with time delay were obtained in the k_p-k_i plane with a fixed k_d for specified GM and PM using the dominant poles method, as reported in [13]. PID controller tuning method for integrating system with time delay was developed in the k_p-k_i plane for a fixed k_d by [15]. After determining the stability region, weighted geometrical center approach was used to determine the required PID controller gains. The results obtained via simulation are good, but the fixed k_d was not optimized. In [16, 17], parametric methods were used to establish the stability region in the k_i-k_d plane for a fixed k_p and k_p-k_i plane for a fixed k_d, and GA was used to determine the optimum controller gains in the established region of the plane.

Despite the huge amount of published works in the literature on PID controller designs for time-delay systems in either k_i-k_d or k_p-k_i design planes, the effects of using either of the two design planes on the performance of the designed system have not been given serious attention in the research community, hence the need for this study. The aim of this study, therefore, is to investigate the effects of PID controller design plane selection on the performance of the designed system.

To achieve the aim of this study, the equations that relate the controller gains, the system parameters, and the time delay together were derived. Using these equations and the frequency obtained from the graph of k_p and k_d against the frequency (ω), the stability boundaries in k_i-k_d and k_p-k_i planes for fixed k_p and k_d respectively was plotted. After the determination of the stability boundaries, the optimum controller gains within the stability region can now be determined. Several optimization methods have been used in the literature but in this study, GA was used because of its heuristic characteristics, powerful searching capabilities [18], and amenability of the method of stability boundary locus (used in this study) to GA. For each of these regions, the optimum controller gains were determined via GA using ITAE as the objective function. Since GA was used as a minimization process in this study, the obtained minimum value of ITAE (Min-ITAE) was plotted against the fixed k_p and k_d for k_i-k_d and k_p-k_i design planes respectively. Based on these graphs, the overall optimum gains were determined. These gains were used together with the system and time-delay transfer functions to generate the system closed-loop unit step response, which yielded the system TDPMs. The TDPMs considered in this study are T_r (rise time), %OS (percentage overshoot), %US (percentage undershoot), T_p (peak time), T_s (settling time), and ess (steady-state error). These performance measures were used for system analysis and characterisation in the plane under consideration. The results show that the k_i-k_d design plane produced a faster response while the k_p-k_i design plane gives a response that closely tracks the input irrespective of the system's type and order. Therefore, the major contribution of the study is to aid the control system designers in the selection of the design plane that will give the best system performance at the beginning of controller design without involving trial and error, given that the system transfer function and design specifications are known.

2. METHODOLOGY

2.1. DETERMINATION OF STABILITY BOUNDARY FOR PID-CONTROLLED SYSTEM

The derivation of system forward and closed-loop transfer functions used in this study is based on the block diagram of the unity feedback control system shown in Fig.1, where *R(s)*, *E(s)*, and *Y(s)* are the reference input, error, and output respectively. Equations (1), (2), and (3) are the adopted expression for the plant, time delay, and controller transfer function respectively.

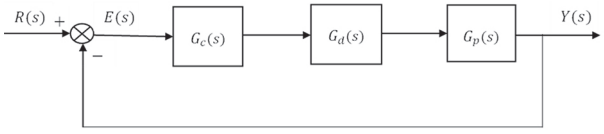


Fig. 1. Block diagram of unity feedback time-delay control system

$$G_p(s) = \frac{N(s)}{D(s)} \tag{1}$$

where *N(s)* and *D(s)* are the plant transfer function numerator and denominator respectively.

$$G_d(s) = e^{-\tau s} \tag{2}$$

where τ is the time delay in sec.

$$G_c(s) = \frac{k_d s^2 + k_p s + k_i}{s} \tag{3}$$

For easy application of D-decomposition method, the numerator and denominator of Equation (1) were broken into their even and odd parts after substituting jω for s as shown in Equation (4). It should be noted that for compactness purpose, the ($-\omega^2$) term has been removed from N_e ($-\omega^2$), N_o ($-\omega^2$), D_e ($-\omega^2$) and D_o ($-\omega^2$) in Equation (4) [19].

$$G_p(j\omega) = \frac{N_e + j\omega N_o}{D_e + j\omega D_o} \tag{4}$$

Finding the stability boundary in k_p, k_i, and k_d space is a three-dimensional problem. For easy controller de-

sign and analysis, it can be reduced to a two-dimensional problem by fixing one of the parameters and finding the stability region in the plane of the remaining two parameters. The overall stability region can be determined using the stability regions in the plane of the two parameters by sweeping over the fixed parameter values. For this study, k_p and k_d were fixed and the stability boundaries in the k_i-k_d and k_p-k_i planes were determined respectively.

The equations and conditions required for determining the stability region in the k_i-k_d plane for a fixed k_p and the stability region in the k_p-k_i plane for a fixed k_d are presented as follows [16, 17, 19, 20]:

2.1.1. Stability region in the k_i-k_d plane for a fixed k_p

For ω=0

$$k_i=0 \tag{5}$$

For ω>0

$$k_p = \frac{\left(\omega^2 N_o D_o + N_e D_e\right)\cos(\omega T)+\omega(N_o D_e - N_e D_o)\sin(\omega T)}{-\left(N_e{}^2+\omega^2 N_o{}^2\right)} \tag{6}$$

$$k_d=\frac{\omega^2(N_o D_e - N_e D_o)\cos(\omega T)-\omega(N_e D_e+\omega^2 N_o D_o)\sin(\omega T)+k_i\left(N_e{}^2+\omega^2 N_o{}^2\right)}{\omega^2\left(N_e{}^2+\omega^2 N_o{}^2\right)} \tag{7}$$

where $\omega = \omega_n$, (n=1,2,……) are the frequencies at which the line of a given value of k_p intercepts the graph of k_p of Equation (6) versus ω, and n is the number of points of intersection or the number of lines obtainable from Equation (7).

The stability boundary in the k_i-k_d plane is formed by the line obtained from Equation (5) and the lines generated from Equation (7) when ω_n is substituted for ω.

2.1.2. Stability region in the k_p-k_i plane for a fixed k_d

For ω=0

$$k_i=0 \tag{8}$$

For $\omega\geq 0$

$$k_p = \frac{\left(\omega^2 N_o D_o + N_e D_e\right)\cos(\omega T)+\omega(N_o D_e - N_e D_o)\sin(\omega T)}{-\left(N_e{}^2+\omega^2 N_o{}^2\right)} \tag{9}$$

$$k_i=\frac{\omega^2(N_o D_e - N_e D_o)\cos(\omega T)-\omega(N_e D_e+\omega^2 N_o D_o)\sin(\omega T)-K_d\omega^2\left(N_e{}^2+\omega^2 N_o{}^2\right)}{-\left(N_e{}^2+\omega^2 N_o{}^2\right)} \tag{10}$$

where ω=[0,ω_c] and ω_c is the frequency at which the line of a given value of k_d in Equation (10) intercepts the graph of k_d in Equation (11) against ω. Equation (11) was obtained from equation (7) when it is assumed that k_i=0.

$$k_d = \frac{\omega^2(N_o D_e - N_e D_o)\cos(\omega T)-\omega\left(N_e D_e+\omega^2 N_o D_o\right)\sin(\omega T)}{\omega^2\left(N_e{}^2+\omega^2 N_o{}^2\right)} \tag{11}$$

The stability boundary in the k_p-k_i plane is formed by the line obtained from Equation (8) and the locus generated by Equations (9) and (10) when ω=[0,ω_c].

2.2. OPTIMIZATION OF PID CONTROLLER PARAMETERS

After the determination of the system convex stability region, the next step is to determine the value of the controller parameters that give the best system performance. The success of the determination and selection of optimum PID controller parameters using GA as an optimization tool is strongly dependent on the selection of appropriate objective function. To optimize the PID controller parameters, different system performance indices have been used as the objective function. ITAE, ISE, MSE, and IAE were used by [21], ISE, ITAE, and IAE were used in [22, 23], while in [24] work, ISE, ITAE, IAE, and their combination were used. The ITAE (see equation (12)) was adopted in this study due to its best selectivity, a good criterion for PID controllers design [25]; and its minimum value can be easily defined as the system parameters are varied [26],

$$ITAE = \int_0^\infty t|e(t)|\,dt \tag{12}$$

where $T_1 \geq T_s$,(T_s is the system settling time in sec.), e(t) is the error and t is the time in sec.

2.3. GA-BASED PID CONTROLLER PARAMETERS OPTIMIZATION

The PID controller gains were optimised using GA by following the steps shown in the flowchart of Fig. 2.

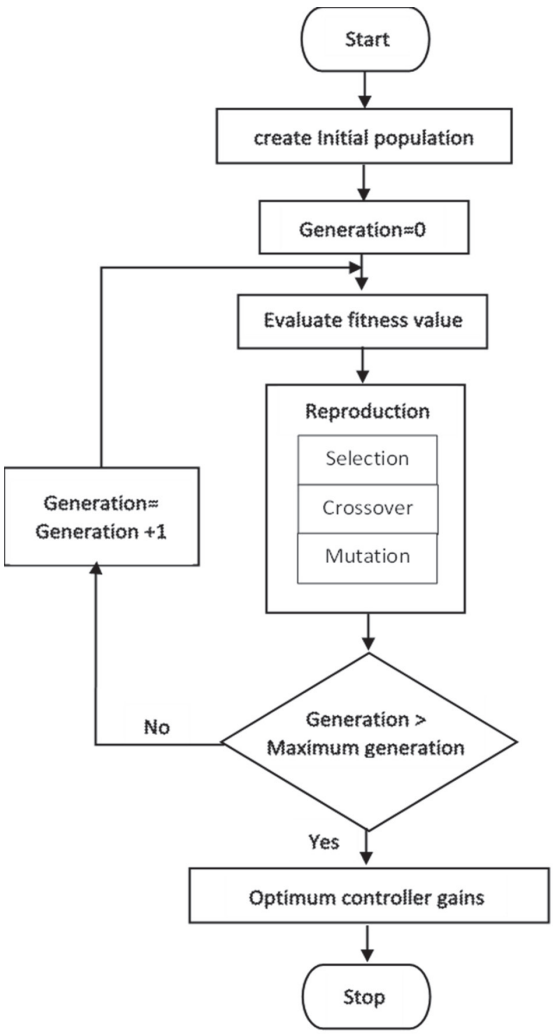


Fig. 2. Flowchart of GA process for optimising PID controller gains [27, 28]

The initial population was created inside the stability region because the stability boundary serves as the constraint for the optimisation problem. To determine the fitness of the generated PID controller gains, the objective function, ITAE in this case was calculated for these gains in parallel. Based on the ranking value of the fitness function for the PID controller gains, half of the population was selected for reproduction. To get the next generation, the crossover was carried out on the already selected feasible solutions (gains). Also, to guide against the production of false optimum controller gains, a few individual solutions in the generation were mutated. The process continues in a loop until when there are no significant changes in the generations which is the condition for termination of the optimisation process. Details on the design of PID controllers for delay-free and time-delay systems using GA can be found in [16, 17] and [29] respectively.

3. SIMULATION EXAMPLES

Three different examples were used to demonstrate the selection of the PID controller design plane. For each of these examples, GA was used in searching for the: (i) optimum k_i and k_d in the k_i - k_d plane for different values of k_p (ii) optimum k_p and k_i in the k_p - k_i plane for different values of k_d. Associated with each set of the optimum gains is a unique Min-ITAE since GA was used as a minimisation process. The optimum fixed k_p gain $k_{p,f,opt}$ for the k_i-k_d plane and $k_{d,f,opt}$ for the k_p - k_i plane is the one that gives Min(Min-ITAE) and can be graphically determined from the graphs of Min-ITAE versus fixed k_p and Min-ITAE versus fixed k_d for k_i-k_d plane and k_p - k_i plane respectively. The overall optimum gains for the k_i-k_d plane are the $k_{p,f,opt}$ and the optimum k_i and k_d denoted by $k_{i,opt}$ and $k_{d,opt}$ respectively, that yielded the Min-ITAE that gives Min(Min-ITAE). Also, the overall optimum gains for the k_p-k_i plane are the $k_{d,f,opt}$ and the optimum k_p and k_i also denoted by $k_{p,opt}$ and $k_{i,opt}$ respectively, that yielded the Min-ITAE that gives Min(Min-ITAE).To provide room for system performance analysis and characterisation in each of these design planes using TDPMs, the system step response was plotted based on the overall optimum controller gains.

3.1. EXAMPLE 1

The design of the PID controller for an integrating second-order time-delay system with the system transfer function given by equation (13) [30] was considered in this example. For this system N_e=1, N_o=0, D_e=-ω^2, and D_o=1.

$$G(s) = \frac{1}{s(s+1)} e^{-s} \tag{13}$$

The ranges of k_d and k_p were first determined by plotting k_d against k_p as shown in Fig. 3 using equations (6) and (11). Also based on these equations, the plots of k_p against ω and k_d against ω are shown in Fig. 4 and 5 for the determination of the relevant frequencies for any given value of k_p and k_d respectively. It can be seen from Fig. 3 that the ranges of k_p and k_d are 0 – 1.717 and -1 – 2.261 respectively.

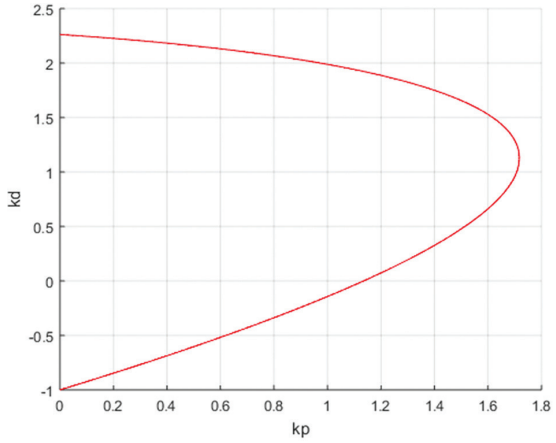


Fig. 3. Plot of k_d against k_p

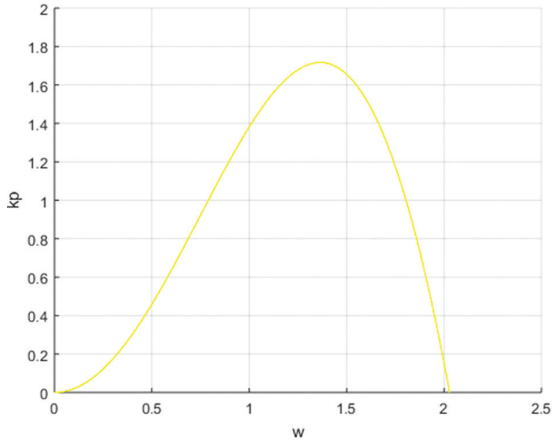


Fig. 4. Plot of k_p against ω

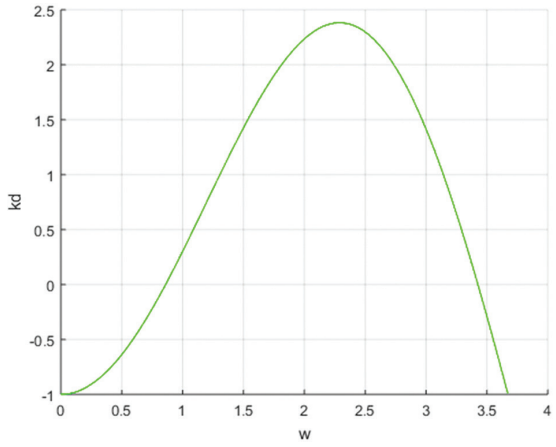


Fig. 5. Plot of k_d against ω

For the case of PID controller design in the k_i-k_d plane, the fixed gains considered based on the range of k_p as presented in Fig. 4 are k_p = 0.2, 0.4, 0.6, 0.8, 1.0, 1.2, 1.4, and 1.6. The detailed design for k_p = 0.2 is presented as follows:

As explained in Subsection 2.1.1., when ω = 0, the stability boundary equation that can be used for generating one of the boundary lines was derived using Equation (5). To generate the equations that can be used to form the

remaining boundary lines, the frequency at the point of intersection of the line of k_p = 0.2 with the plot of k_p versus ω in Fig. 4 was determined. From Fig. 4, the points of intersection are two and their corresponding frequencies are ω_1=0.322 rad and ω_2= 1.991 rad. Substituting ω_1 and ω_2 in Equation (7) yielded the remaining two equations required for generating the stability boundary's lines. The resulting stability boundary equations are:

$$k_i=0 \quad (14)$$

$$k_{d1}=9.6447k_i-0.8467 \quad (15)$$

$$k_{d2}=0.2523k_i+2.2257 \quad (16)$$

Based on Equations (14) – (16), and the similar equations obtained for values of remaining fixed k_p following, the steps used for k_p = 0.2, the system stability boundary in the k_i-k_d plane shown in Fig. 6 and 7 were obtained for when k_p = 0.2 and for the complete set of fixed k_p, respectively.

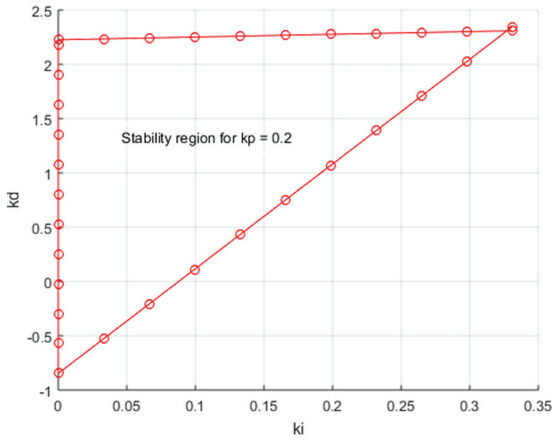


Fig. 6. The stability boundary in the k_i-k_d plane for k_p = 0.2

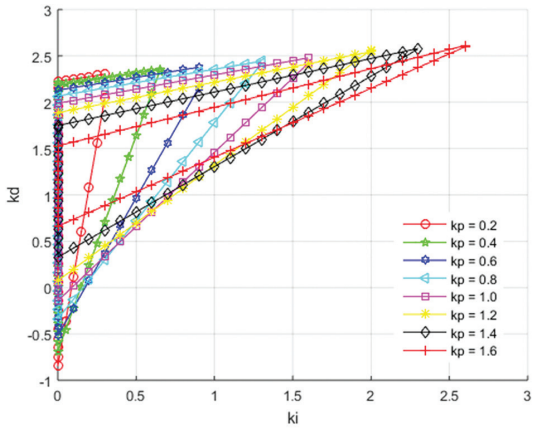


Fig. 7. The stability boundary in the k_i-k_d plane for the set of fixed k_p

The constraints needed for the GA were formed using Equations (14) – (16) and the stability boundary equations for other fixed k_p considered following the same steps. After the formation of the constraints, the GA MATLAB code was written according to the flowchart in Fig. 2. The code was run for each scenario to determine the optimum k_i and k_d, and the corresponding Min-ITAE for each of the fixed k_p. The $k_{p,f,opt}$ and Min (Min-ITAE) were obtained from the plot of Min-ITAE versus fixed k_p. The required TDPMs were obtained from the unit step response plotted using the determined overall optimum controller gains.

Taking the case of PID controller design in the k_p-k_i plane the fixed gains considered based on the range of k_d as shown in Fig. 5 are k_d = -0.5, 0, 0.5, 1.0, and 1.5. The design analysis for k_d = -0.5 is presented as follows:

As discussed in Subsection 2.1.2., when ω = 0, the stability boundary equation that was used for generating one of the boundaries was derived using Equation (8). The frequency at the first point of intersection of the line of k_d = -0.5 with the plot of k_d versus ω in Fig. 5 was determined to generate the equation that can be used to form the remaining boundary locus. From Fig. 5, the frequency at the said points of intersection, ω_c=0.5917 rad. By generating the straight line defined by Equation (8) and plotting k_i against k_p using Equations (9) and (10) with ω=[0, ω_c] the stability boundary of Fig. 8 was obtained for k_d = -0.5.

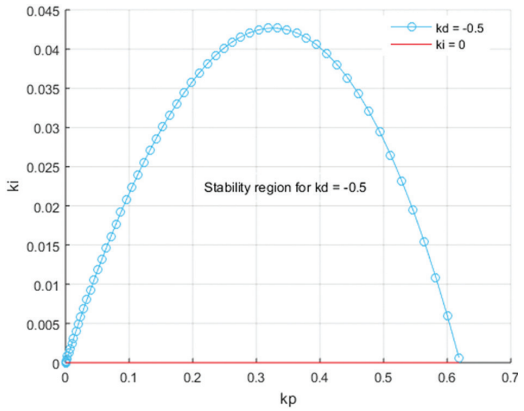


Fig. 8. The stability boundary in the k_p-k_i plane for k_d = -0.5

The same steps used for k_d = -0.5 was adopted for other fixed k_d scenarios and the stability boundaries of Fig. 9 were obtained. The stability boundary equations for k_d = -0.5 case were obtained from Equation (8) and by curve fitting the stability locus of Fig. 8 and are presented in Equations (17) and (18) respectively.

$$k_i=0 \quad (17)$$

$$k_i= -0.061k_p^4-0.069k_p^3-0.34k_p^2+0.25k_p-3.6\times10^{-6} \quad (18)$$

For other fixed k_d values, the stability boundary equations like (17) and (18) were also generated following the steps used for k_d = -0.5 case.

The constraints needed for the GA were formed using Equations (17) and (18) and the stability boundary equations for other fixed k_d were considered. After the formation of constraints, the GA MATLAB code was written according to the flowchart in Fig. 2. The code was run for each scenario to determine the optimum k_p and k_i,

and the corresponding Min-ITAE for each of the fixed k_d. $k_{d,f,opt}$ and Min (Min-ITAE) were obtained from the plot of Min-ITAE versus fixed k_d. Also, the required TDPMs were obtained from the unit step response plotted using the determined overall optimum controller gains.

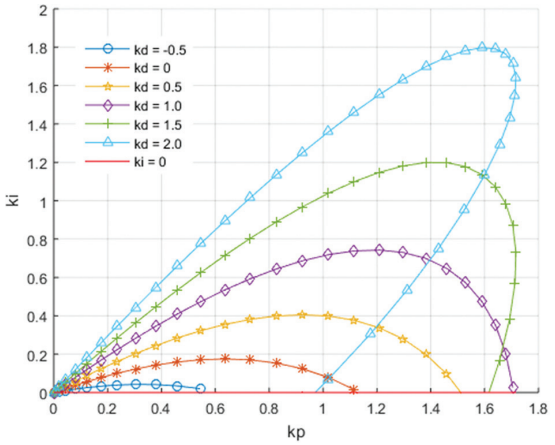


Fig. 9. The stability boundary in the k_p-k_i plane for the set of fixed k_d

3.2. EXAMPLE 2

The design of a PID controller for an integrating second-order time-delay system with left-hand side zero whose transfer function is shown in Equation (19) [15] was considered in this example. For this system N_e=0.6, N_o=-0.18, D_e= -w^2, and D_o=1

$$G(s) = \frac{0.6(-0.3s+1)e^{-0.2s}}{s(s+1)} \tag{19}$$

The method employed for the design of the PID controller in Example 1 was adopted in this example. As a result, the PID controller design in the k_i-k_d plane in this example was considered for the following fixed k_p values:1, 2, 3, 4, 5, and 6. The stability boundary for the considered values of fixed k_p is shown in Fig. 10. The required TDPMs were obtained from the unit step response plotted using the determined overall optimum controller gains.

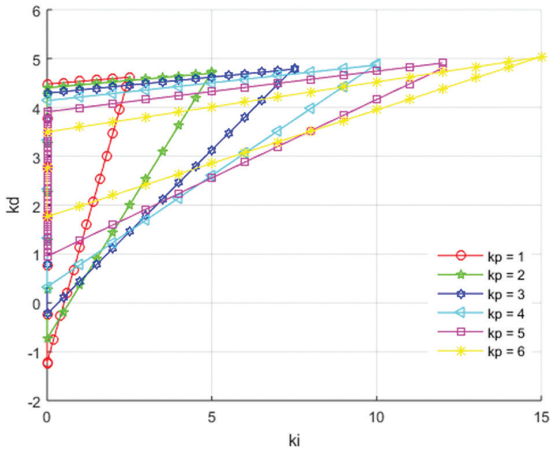


Fig. 10. The stability boundary in the k_i-k_d plane for the set of fixed k_p

To design the PID controller in the k_p-k_i plane for this example the considered range of fixed k_d are k_d = -1, 0, 1, 2, 3, and 4, and the stability boundaries of Fig. 11 were obtained for these fixed values of k_d. The required TDPMs were obtained from the unit step response plotted using the determined overall optimum controller gains.

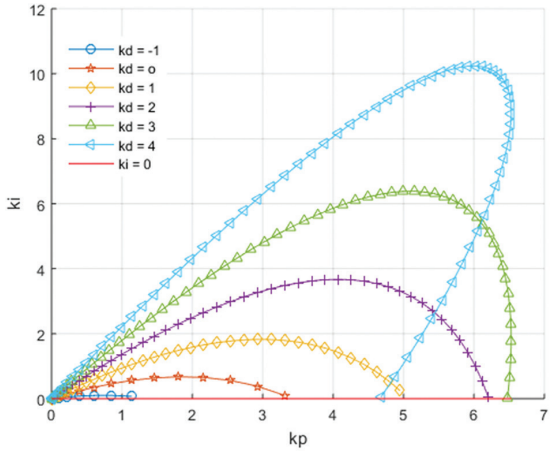


Fig. 11. The stability boundary in the k_p-k_i plane for the set of fixed k_d

3.3. EXAMPLE 3

The design of a PID controller for a second-order time-delay system with left-hand side zero whose transfer function is shown in Equation (20) [19] was considered in this example. For this system N_e=1, N_o=-0.5, D_e=1-2w^2, and D_o=3.

$$G(s) = \frac{(-0.5s+1)e^{-0.6s}}{(s+1)(2s+1)} \tag{20}$$

In this example, the method of PID controller design used in Example 1 was adopted. Therefore, for the k_i-k_d plane the considered fixed k_p values are: k_p = -0.5, 0, 0.5, 1.0, 1.5, 2.0, 2.5, 3.0, and 3.5. For these k_p values, the obtained stability boundary is shown in Fig. 12. The required TDPMs were obtained from the unit step response plotted using the determined overall optimum controller gains.

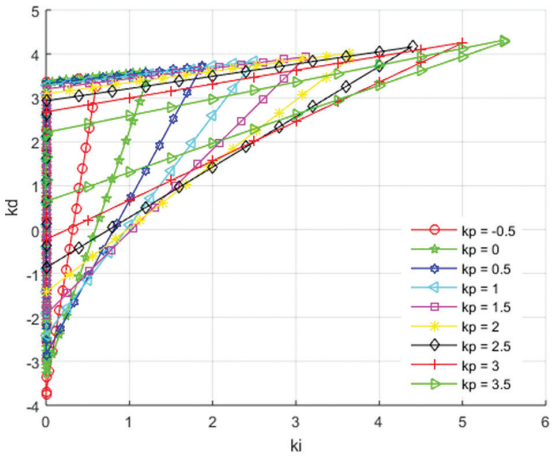


Fig. 12. The stability boundary in the k_i-k_d plane for the set of fixed k_p

For the design of PID controller in the k_p-k_i plane the fixed values of k_d considered are: k_d = -3, -2, -1, 0, 1, 2 and 3. Based on these values of k_d the stability boundaries plot of Fig. 13 was obtained. The required TDPMs were obtained from the unit step response obtained using the determined overall optimum controller gains.

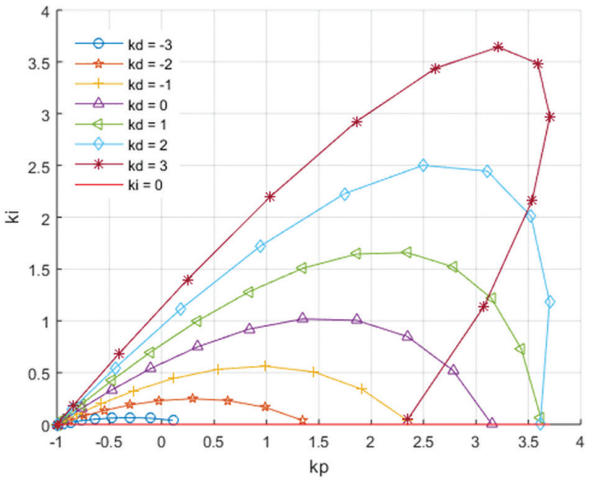


Fig. 13. The stability boundary in the k_p-k_i plane for the set of fixed k_d

4. RESULTS AND DISCUSSIONS

The graphs of Min-ITAE versus the fixed k_p and fixed k_d for the k_i-k_d design plane and k_p-k_i design plane for Examples 1, 2, and 3 are shown in Fig. 14, 15, and 16 respectively. It can be seen from these Figures that Min-ITAE decreases from the initial value to the minimum value and then increases to a final value. Based on the minimum of Min-ITAE and the corresponding fixed k_p and fixed k_d obtained, $k_{p,f,opt}$, $k_{d,f,opt}$ and Min (Min-ITAE) were determined and presented in Table 1. Using the value of $k_{p,f,opt}$, $k_{d,f,opt}$ and the corresponding value of Min (Min-ITAE) the value of $k_{p,opt}$, $k_{i,opt}$ and $k_{d,opt}$ were obtained from the GA optimisation results and presented in Table 1.

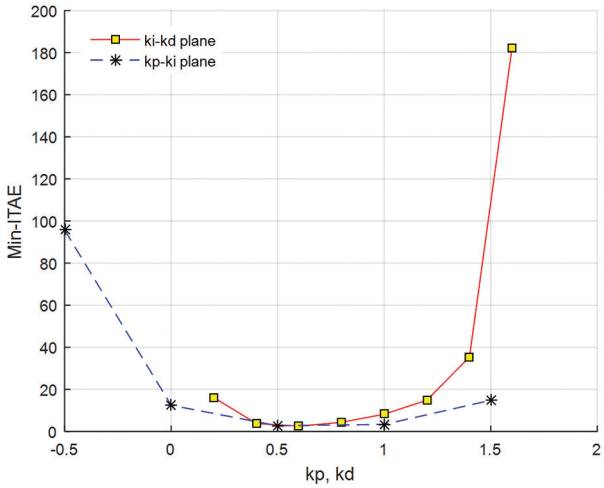


Fig. 14. Plot of Min-ITAE versus the values of fixed k_p, k_d for Example 1

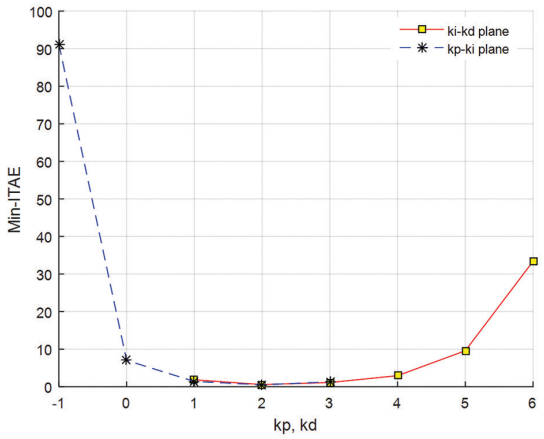


Fig. 15. Plot of Min-ITAE versus the values of fixed k_p, k_d for Example 2

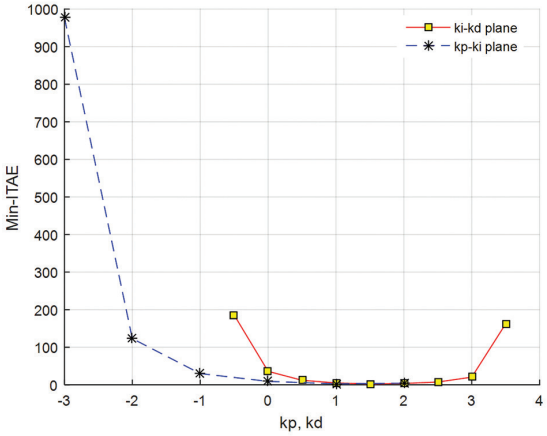


Fig. 16. Plot of Min-ITAE versus the values of fixed k_p, k_d for Example 3

Using the open-loop transfer functions of Equations (13), (19), and (20) and the overall optimum gains of Table 1, the system closed-loop unit step responses of Fig. 17, 18, and 19 were plotted for Examples 1, 2, and 3 respectively. From Fig. 17, 18, and 19, it is obvious that the systems' response never tracks the unit step input instead, 0.5, 0.375, and 0 were tracked respectively, therefore, the need for a controller for these systems.

The TDPMs values obtained from the step response of Fig. 17 to 19 for the controlled systems under the two design planes are presented in Table 2. It is obvious from Table 2 that the resulting system from the two design planes for the 3 examples is of good steady state because *ess*=0. Considering the transient state status, the response of the k_i-k_d design plane has higher swiftness for all the examples because of low T_r and T_p compared to the k_p-k_i design plane. On the other hand, the degree of similarity between the response of the system designed in the k_p-k_i plane with the unit step input for all examples is high compared with that of the k_i-k_d plane because of the lower value of %***OS*** and T_s associated with the former. It can also be seen from Table 2, for Examples 2 and 3, there is an associated %*US* because of the right-hand side zero. But the %*US* for the k_i-k_d plane is higher than that of the k_p-k_i plane.

Table 1. Optimum controller gains for different scenarios

Example	Plane	$k_{p,f,opt}$	$k_{d,f,opt}$	$k_{p,opt}$	$k_{i,opt}$	$k_{d,opt}$	Min (Min-ITAE)
1	k_i-k_d	0.6			-1.133×10^{-11}	0.712	2.538
	k_p-k_i		0.5	0.470	3.338×10^{-07}		2.703
2	k_i-k_d	2.000			-1.808×10^{-09}	2.278	0.5762
	k_p-k_i		2.000	1.935	2.013×10^{-12}		0.5184
3	k_i-k_d	1.5			0.485	1.113	2.398
	k_p-k_i		1.000	1.410	0.460		2.564

Table 2. TDPMs for different examples and planes

Example	Plane	T_r (sec.)	*%OS*	*%US*	T_p (sec.)	T_s (sec.)	*ess*
1	k_i-k_d	1.293	10.222	0	3.522	6.720	0.000
	k_p-k_i	2.067	1.092	0	5.083	3.935	0.000
2	k_i-k_d	0.437	14.048	41.009	1.421	2.940	0.000
	k_p-k_i	0.549	7.926	36.000	1.618	2.136	0.000
3	k_i-k_d	1.485	1.874	27.812	3.985	3.202	0.000
	k_p-k_i	1.736	0.948	25.000	4.570	3.066	0.000

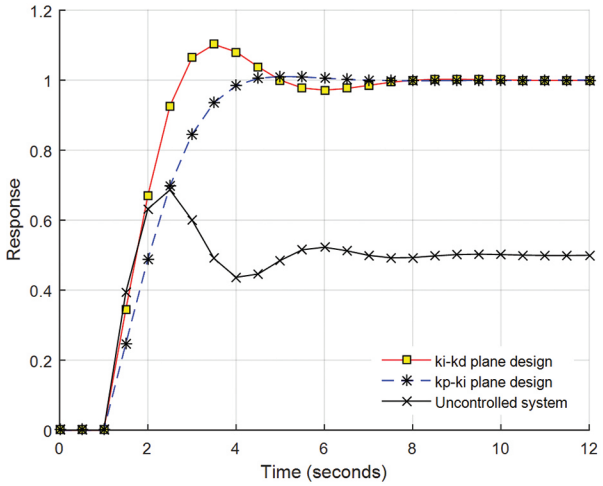


Fig. 17. The unit step response using overall optimum gains for Example 1

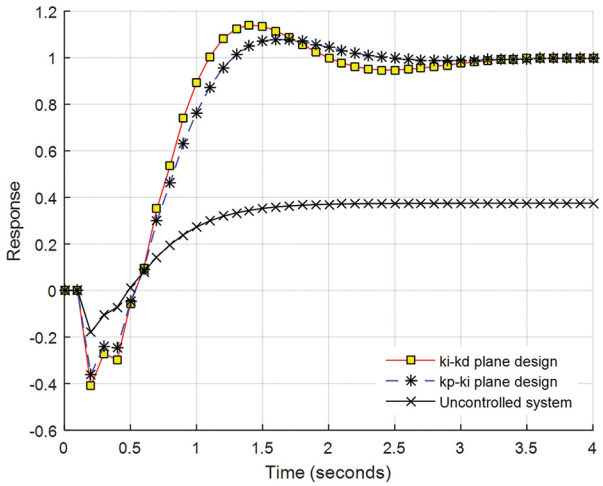


Fig. 18. The unit step response using overall optimum gains for Example 2

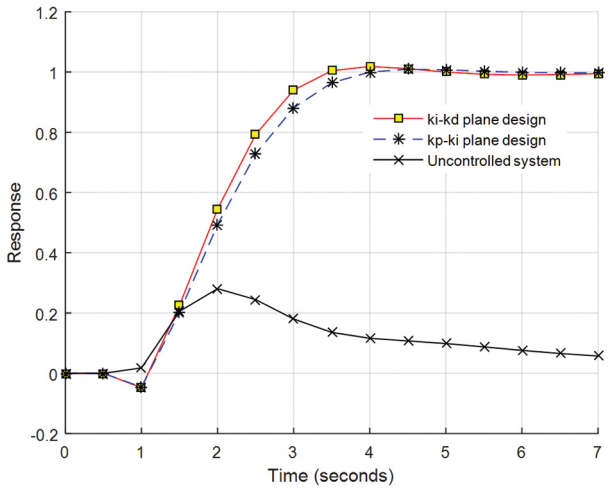


Fig. 19. The unit step response using overall optimum gains for Example 3.

It was observed that irrespective of the type and order of the system to be controlled, the k_i-k_d design plane yielded a faster response while the k_p-k_i design plane yielded a response that closely tracks the input. From these results, the system design plane can be decided by control system designers since system design specifications depend on its applications.

5. CONCLUSIONS

The parametric effects of design plane selection on the performance of PID-controlled time-delay systems were presented in this study. System stability region in the k_i-k_d and k_p-k_i planes were determined for a fixed value of k_p and k_d respectively. The study concluded that systems required controllers because their closed-loop response can never track the reference input. The controllers were designed by optimising the controller

gains using a combination of GA and graphical methods to get the overall optimum gains.

It was also concluded, that irrespective of the type and order of the system, the k_i-k_d design plane yielded a system with a faster response while the k_p-k_i design plane yielded a system response that closely track the input. Also, for a system with right-hand-side zero the k_i-k_d design plane results in a system response with a higher $\%US$ compared to the k_p-k_i design plane.

The study contributes to aiding the control system designers to select the design plane which gives the best performance right from the start of controller design without using a trial and error approach, once the system transfer function and design specifications are known.

6. REFERENCES

[1] S. Testouri, K. Saadaoui, M. Benrejeb, "Robust stabilization of a class of uncertain systems with time delay", IFAC Proceedings Volumes, Vol. 43, No. 8, 2010, pp. 553-557.

[2] L. Barhoumi, I. Saidi, D. Soudani, "Graphical method for obtaining PID parameters for systems with time delay", International Journal of Computer Science and Network Security, Vol. 19, No. 7, 2019, pp. 31-37.

[3] A. Yüce, N. Tan, D. P. Atherton, "Fractional order pi controller design for time delay systems", IFAC-PapersOnLine, Vol. 49, No. 10, 2016, pp. 94-99.

[4] J. P. Richard, "Time-delay systems: an overview of some recent advances and open problems", Automatica, Vol. 39, No. 10, 2003, pp. 1667-1694.

[5] L. M. Eriksson, M. Johansson, "PID controller tuning rules for varying time-delay systems", Proceedings of the American Control Conference, New York, NY, USA, 11-13 July 2007, pp. 619-625.

[6] A. Gupta, S. Goindi, G. Singh, H. Saini, R. Kumar, "Optimal design of PID controllers for time delay systems using genetic algorithm and simulated annealing", Proceedings of International Conference on Innovative Mechanisms for Industry Applications, Bengaluru, India, 21-23 February 2017, pp. 66 -69.

[7] N. Hohenbichler, "All stabilizing PID controllers for time delay systems", Automatica, Vol. 45, No. 11, 2009, pp. 2678-2684.

[8] F. A. Salem, "New Efficient Model-Based PID Design Method", European scientific journal, Vol. 9, No. 15, 2013, pp. 181-199.

[9] L. Ou, W. Zhang, D. Gu, "Sets of stabilising PID controllers for second-order integrating processes with time delay", IEE Proceedings-Control Theory and Applications, Vol. 153, No. 5, 2006, pp. 607-614.

[10] H. Efheij, A. Albagul, "Comparison of PID and artificial neural network controller in online of real-time industrial temperature process control system", Proceedings of the IEEE 1st International Maghreb Meeting of the Conference on Sciences and Techniques of Automatic Control and Computer Engineering MI-STA, Tripoli, Libya, 25-27 May 2021, pp. 110-115.

[11] N. B. Hassen, K. Saadaoui, M. Benrejeb, "Stabilizing lead lag controllers for time delay systems", Recent Advances on Electroscience and Computers, Vol. 1, No. 2, 2015, pp. 106-109.

[12] C. Onat, "A new design method for PI–PD control of unstable processes with dead time", ISA Transactions, Vol. 84, 2019, pp. 69-81.

[13] S. Srivastava, V. S. Pandit, "A PI/PID controller for time delay systems with desired closed loop time response and guaranteed gain and phase margins", Journal of Process Control, Vol. 37, 2016, pp. 70-77.

[14] S. E. Hamamci, "An algorithm for stabilization of fractional-order time delay systems using fractional-order PID controllers", IEEE Transactions on Automatic Control, Vol. 52, No. 10, 2007, pp. 1964-1969.

[15] M. M. Ozyetkin, C. Onat, N. Tan, "PID tuning method for integrating processes having time delay and inverse response", IFAC PapersOnLine, Vol. 51, No. 4, 2018, pp. 274-279.

[16] K. Saadaoui, S. Elmadssia, M. Benrejeb, "Stabilizing PID controllers for a class of time delay systems", PID Controller Design Approaches-Theory, Tuning and Application to Frontier Areas, IntechOpen, 2012, pp. 141-158.

[17] K. Saadaoui, A. Moussa, M. Benrejeb, "PID controller design for time delay systems using genetic algorithms", The Mediterranean Journal of Measurement and Control, Vol. 5, No. 1, 2009, pp. 31-36.

[18] A.Y. Jaen-Cuellar, R. de J. Romero-Troncoso, L. Morales-Velazquez, R.A. Osornio-Rios, "PID-controller tuning optimization with genetic algorithms in servo systems", International Journal of Advanced Robotic Systems, Vol. 10, No. 9, 2013 pp. 324-337.

[19] N. Tan, "Computation of stabilizing PI and PID controllers for processes with time delay", ISA Transactions, Vol. 44, No. 2, 2005, pp. 213-223.

[20] S. Sujoldzic, J. M. Watkins, "Stabilization of an arbitrary order transfer function with time delay using PI and PD controllers", Proceedings of the American Control Conference, Minneapolis, MN, USA, 14-16 June 2006, pp. 2427-2432.

[21] J. Zhao, M. Xi, "Self-Tuning of PID parameters based on adaptive genetic algorithm", in IOP Conference Series: Materials Science and Engineering, Vol. 782, No. 4, 2020, pp. 1-8.

[22] M. Angelova, T. Pencheva, "Tuning genetic algorithm parameters to improve convergence time", International Journal of Chemical Engineering, 2011, pp. 1-7.

[23] A. A. Chlaihawi, "Genetic algorithm error criteria as applied to PID controller DC-DC buck converter parameters: an investigation", IOP Conference Series: Materials Science and Engineering, Vol. 671, No. 1, 2020, pp. 1-10.

[24] S. Mahfoud, A. Derouich, N. EL Ouanjli, M. EL Mahfoud, M. Taoussi, "A New Strategy-based PID controller optimized by genetic algorithm for DTC of the doubly fed induction motor", Systems, Vol. 9, No. 2, 2021, pp. 1-18.

[25] O. Guenounou, B. Dahhou, B. Athmani, "Optimal design of PID controller by Multi-objective genetic algorithms", Proceedings of the International Conference on Computer Related Knowledge, Sousse, Tunisia, July 2012, pp. 1-6.

[26] K. M. Hussain, R. A. R. Zepherin, M. S. Kumar, S. M. Giriraj Kumar, "Comparison of PID controller tuning methods with genetic algorithm for FOPTD system", International Journal of Engineering Research and Applications, Vol. 4, No. 2, 2014, pp. 308-314.

[27] E. Gouthami, M. A. Rani, "Modeling of an adaptive controller for an aircraft roll control system using PID, fuzzy-PID and genetic algorithm", Journal of Electronics and Communication Engineering, Vol. 11, No. 1, 2016, pp. 15-24.

[28] S. Pareek, M. Kishnani, R. Gupta, "Comparative Analysis of GA-PID Controller for different performance indices", Proceedings of STEER 2014: ICIEEC, 2014, pp. 48-53.

[29] A. A. M. Zahir, S. S. N. Alhady, W.A.F.W. Othman, A. A. A. Wahab, M. F. Ahmad, "Objective functions modification of GA optimized PID controller for brushed DC motor", International Journal of Electrical and Computer Engineering, Vol. 10, No. 3, 2020, pp. 2426-2433.

[30] S. Atiç, E. Cokmez, F. Peker, I. Kaya, "PID controller design for controlling integrating processes with dead time using generalized stability boundary locus", IFAC-Papers Online, Vol. 51, No. 4, 2018, pp. 924-929.

Cost Prediction for Roads Construction using Machine Learning Models

Original Scientific Paper

Yasamin Ghadbhan Abed
University of Diyala, College of Science, Department of Computer Science
Diyala, Iraq
scicompms2133@uodiyala.edu.iq

Taha Mohammed Hasan
University of Diyala, College of Science, Department of Computer Science
Diyala, Iraq
dr.tahamh@uodiyala.edu.iq

Raquim Nihad Zehawi
University of Diyala, College of Engineering, Department of Highway and Airport Engineering
Diyala, Iraq
raquim_zehawi@uodiyala.edu.iq

Abstract *– Predicting conceptual costs is among the essential criteria in project decision-making at the early stages of civil engineering disciplines. The cost estimation model availability that may help in the early stages of a project could be incredibly advantageous in respect of cost alternatives and more extraordinary cost-effective solutions periodically. There is a lack of case datasets. Most of the proposed dataset was inefficient. This study offers a new data set that includes the elements of road construction and economic advantages in the year of project construction. Real project data for rural roads in the State of Iraq / Diyala Governorate for the years 2012 to 2021 have use to train a predictive model with a high rate of accuracy based on machine learning (ML) methods. Ridge and Least Absolute Shrinkage and Selection Operator (LASSO) Regressions, K Nearest Neighbors (k-NN), and Random Forest (RF) algorithms have employ to create models for estimating road construction costs based on real-world data. The Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE), and R-squared (R2) coefficient of determination are utilize to assess the models' performance. The analysis indicated that the RR is the best model for road construction costs, with results R2 = 1.0, MAPE =0.00, and RMSE=0.00. The results showed that the cost estimates were accurate and aligned with the project bids.*

Keywords*: Construction, Roads, Cost estimation, Machine learning, Ridge regression*

1. INTRODUCTION

Cost estimation is one of the essential concerns in the early phases of a construction project's life cycle. The cost of a building project is widely acknowledged as critical contract data, yet it is frequently miscalculated. Contracted prices can result in various issues, resulting in additional expenditures throughout the project's implementation. The primary and most common problem in many building projects is a cost overrun. [1]

If the inaccuracy of the initial cost estimates was merely due to inadequate information and inherent forecasting challenges, as those in charge of project estimates indicated, the inaccuracies might be expected to be random. [2]

Given the detrimental consequence on profitability and public funds, extra expenses are a key source of worry for both the private and public sectors. Since the contract is closed by a predetermined amount of the bid, road projects are expected to exceed the budget. The additional expenditures of projects have been detected as a result of deletions, errors, and contract changes. [3]

TThe necessity for a precise preliminary estimate has prompted study into constructing models relying on machine learning (ML) algorithms to predict the initial assessment of road, building, bridge, or other construction projects. [4]

2. LITERATURE SURVEY

The creation and use of methodologies for the cost estimation of road projects have been active research areas over the past few decades. Numerous studies on

estimating construction costs using neural networks, regression, or stochastic methods have been published in the last 20 years.[5]

The dataset is a crucial vector in developing a predictive model for cost estimation. Some authors used a small dataset to analyze the effect of variables on cost using ML techniques like [4]. Cost models were developed utilizing 50 sets of data gathered from road projects finished in South Western Nigeria between 2010 and 2015. They employed linear and multiple regression to anticipate the preliminary estimate of road projects for seven primary construction activities. Based on data collected from the Brazilian National Department of Transport Infrastructure (DNIT), fourteen highway projects in Brazil have been utilized to establish a more precise estimation technique for the construction of highway projects utilizing Artificial Neural Networks (ANNs) Barros, Marcy, and Carvalho. [6] The inputs have been the most impactful factors in the road project estimation costs, and the output was the actual cost value of the work. An average cost estimation accuracy of 99% was accomplished. Furthermore, Tijanić et al. [7] examined the performance of several types of artificial neural networks (multilayer perceptron MLP, generalized regression neural network (GRNN), GRNN, radial basis function neural network RBFNN) for estimating road construction costs in Croatia, using a dataset of 57 road sections. The GRNN had the greatest precision, with a Mean Absolute Percentage Error (MAPE) of 13% and a coefficient of determination of 0.95. According to Mahalakshmi and Rajasekaran [8], a multi perceptron network with a backpropagation algorithm is competent of accurately forecasting highway construction costs. The National Highway Authority of India (NHAI) provided a dataset of 52 projects. Subsequently, AL-Zwainy and Aidan[9] provided multi-layer perceptron training that utilized the backpropagation algorithm to predict construction costs of highways in Iraq. The dataset was 150 past highway data from the republic of Iraq, and it was not published. The ANNs model was able to forecast the cost of structural work for a highway project with a high degree of accuracy (93.19%) and a high coefficient of correlation (R) of 90.026%. Peško et al., [10] analyzed support vector machines (SVMs) and ANNs using SVM have shown higher precision when estimating costs, with MAPE of 7.06% contrast to the extremely accurate ANNs, which have attained a 25.38% precision. The dataset was 166 projects. Moreover, Suneja et al., [11] focused on developing a cost estimation model for Transportation Infrastructure Projects based on Reality by Neural Network to discover the connection between multiple variables of the project and their cost. The dataset was 124 road projects in Gujarat Region. In Poland, a number of completed bridge construction projects were collected by [12] to build an SVM-based regression model to predict bridge construction costs with accuracy appropriate for the early stage of projects. The model was capable of providing an early estimate with satisfactory accuracy reach to 0.98 for the correlation coefficient of real-life bridge construction costs, but the dataset was not clear. [13] used multiple regression techniques for develop early cost estimating models for road construction projects , based on 131 sets of data collected in the West Bank in Palestine. R2 for the developed models was varying from 0.92 to 0.98 which indicates that the predicted values from estimated models fit with the real-life data.

According to everything mentioned above, the esearchers focused on determining the most affected variable on estimation cost. Accuracy was not considered during the model design since the work is more analytical than artificial intelligence (AI). Different datasets have been used in every research, some very small, and all the datasets were not published to compare the results. This study presented and published a real dataset for road construction in building a predictive model for a very high accuracy cost based on ML techniques.

There is a continual need for competent computing methods in this application field due to economic and environmental constraints and their linkages in road construction development. In recent years, AI-based ML algorithms have been verified to be superior to traditional methods for making such forecasts in a variety of infrastructure development projects. Machine learning approaches aim to predict, explain, and discover correlations and patterns between variables [14].

The adverse effects of biased cost modeling in the construction industry are significant since such modeling may drastically reduce project costs by underestimating or overestimating costs. As a result, engineers and managers need this information to swiftly assess alternative project options' feasibility, performance, and profitability [5]. Therefore, a data configuration is proposed that contains a set of material and financial variables for actual road construction and economic variables affecting road construction costs [15]. Historical rural road construction projects in Diyala Governorate were chosen for estimating and modeling the total construction costs.

This study's contribution is to design a new and more realistic road construction cost estimating model that incorporates advanced ML concepts, economic data, and indices. The proposed approach compares four ML algorithms. Ridge and Least Absolute Shrinkage and Selection Operator (LASSO) Regressions, K Nearest Neighbors (k-NN), and Random Forest (RF) algorithms. The technique successfully assisted stakeholders in the early stages of a construction project who were responsible for estimating and managing construction costs to accomplish more precise findings from previous situations.

The remainder of the work is organized as follows: Section 2 describes the case study. Section 3 explains the modeling methodology. The experimental results of the approach and a commentary on the findings are presented in Section 4. Finally, Section 5 concludes the

report with closing remarks and recommendations for further research.

3. METHODS AND MATERIAL

Artificial intelligence (AI) predicts or calculates the cost of construction-based materials or construction datasets. The machine learning (ML) approach is a primary field concerning AI for predicting classes or targets with accurate results. ML was approached into the superior and un-superior methods. This section explains the theoretical concept concerning ML methods in subsection 2.1. Also, the materials or datasets described in subsection 2.2 are collected and proposed for training and testing ML algorithms based on evaluation metrics.

3.1 METHODS

ML algorithms allow for more complex cost prediction models. They learn from input variables and provide data-driven predictions on output variables instead of static prediction models resembling those used in time series analysis. In addition, explanatory variables (also referred to as features in the ML context) improve the capacity of a machine learning model to detect variance and deliver a more accurate prediction [15][16]. Four ML algorithms have been used and analyzed. Ridge and Least Absolute Shrinkage and Selection Operator (LASSO) Regressions, K Nearest Neighbors (k-NN), and Random Forest (RF) algorithms.

3.1.1 K Nearest Neighbors (k-NN) Regression

The computation of the k-NN is a well-known and valuable supervised learning method that employs the concept of similarity to predict a target output (for example, a class label) for a query object or sample. The k-NN method insinuates the intended output of new objects in the feature space of a training set depending on the outcomes of the nearest samples or the outcome of many nearest objects. [17] The k-NN regression is a technique for gradual learning based on occurrences. A nonparametric regression accelerates the training phase since it imposes no assumptions about data distribution. It learns complex target functions rapidly as well as without losing any data. K observations with x_i in close proximity are considered for a particular input x of training data, and the average of the responses of those K independent variables produces $y\hat{}$

$$y\hat{}(x) = \frac{1}{k}\sum_{x_i \in N_{k(x)}} y_i \quad (1)$$

where $N_{k(x)}$ illustrates K closest points in the neighborhood of x. Various distance metrics are used to determine how close two points are, but Euclidean distance is the most often used [17] [18].

3.1.2 Ridge Regression

During the 1970s and 1980s, a newly developed approach for calculating multiple linear regression coefficients called Ridge Regression (RR) was one of the most intriguing research subjects [19]. Ridge Regression is a well-known parameter estimation method for dealing with the collinearity issue that commonly occurs in multiple linear regression [20]. The RR is a tool for assessing multicollinearity data from multiple regression models. The RR is also crucial for analyzing multicollinearity in multiple regression data. Least-squares evaluations are impartial when multicollinearity occurs, but their modifications are greater. Thus, they may be far from their true value. By adding a bias grade to the regression evaluations, RR decreases standard errors. It is expected that, as a result, more consistent evaluations will be available. In addition, when the loss function is the linear least-squares function, and the data is regularized using the L2-norm, the RR model may be utilized to solve a regression problem. The strength of the regularization has to be a positive float. Regularization enhances the conditioning of the problem and lowers the estimated variance [21]. This strategy was initially presented to handle the multicollinearity problem by Hoerl and Kennard (1970) [22]. They proposed that a small positive number be added to the diagonal elements of the $X'X$ matrix, yielding the estimators Eq. 3:

$$\beta = (X^t X)^{-1} X^t Y \quad (2)$$

$$\beta\hat{} = (\ X\hat{}' X + kI\_p\ )\hat{}(-1)\ X\hat{}'\ y\ , k \geq 0 \quad (3)$$

This is referred to as a ridge regression estimator, and the constant k ($k \geq 0$) is referred to as a "biased" or " ridge" parameter that must be estimated with real data.

Algorithm (1): Ridge Regression Algorithm

Input: preprocessed data

Output: Road Cost

Begin

Step 1: Load the Data

Step 2: Creating a New Train and Validation Datasets (train_test_split)

Step 3: Classifying Predictors and Target, Classifying Independent and Dependent Features.

Step 4: Evaluating The Model With the R-squared , MAPE , and RMSE

Step 5: Building the Ridge Regressor (Initializing the Ridge Regressor with alpha =1.0)

Step 6: Fitting the Training data to the Ridge regressor

Step 7: Predicting for X_test

Step 8: calculate the R-squared of the model on the training data

Step 9: Calculate the MAPE of the model on the training data

Step 10: Calculate the MSE and RMSR of the model on the training data

Return Cost value

End

3.1.3 Least Absolute Shrinkage and Selection Operator (LASSO) Algorithm

The LASSO is a popular regression approach that achieves a sparse answer using a l_1 penalty. The LASSO is also known as the basis pursuit in the signal processing literature. For example, generalized linear models, as well as Cox's proportional hazard models for survival data, have been widely employed. LASSO is uninterested in highly linked predictors, preferring to pick one and ignoring the rest. Many coefficients should be close to zero, while a small fraction should be larger and nonzero, according to the LASSO penalty. It's a linear model with a regularization term added to it mathematically, as in Eq. 4. The function to minimize as an objective function is as follows: [23].

$$\min_{w} \frac{1}{2n_{samples}} \left\| x_{w-y} \right\|_2^{2^{i=1}} + \alpha \|w\|_1 \quad (4)$$

The LASSO estimate, therefore, solves the least-squares penalty minimization with $\alpha||w||_1$ added, in which α refers to a constant and $||w||_1$ refers to the l_1 norm of the coefficient vector [24].

2.1.4 Random Forest (RF) algorithm

RF is a more advanced classification and regression decision tree approach. It is also a member of the learner ensemble. Because of its simple structure, a decision tree is a simple method to employ. Unfortunately, because of the enormous variance, it is unstable. A random forest appears to solve the problem. RF is a process for creating numerous independent decision trees with varied sets of samples at each node and averaging the scores of each decision tree as the final score to achieve a more precise outcome [25]. The algorithm constructs a forest with a number of decision trees during training. A set of decision nodes divides a tree into its many branches until it achieves the termination point (the leaf), which is the decision tree's prediction. Each decision node is dependent on if the value of input features is higher than or equivalent to a threshold value. Every forest tree is presented in a subtly distinct approach to mimic a model. The resulting prediction is obtained by averaging each of the tree forecasts [26].

3.2 MATERIAL

In this study, the dataset contain about 1660 project encompasses at least one construction item, and every project has 24 features. The data for the construction items group and the economic data group was separated, as shown in Table 1.

Table 1. Construction data group

Feature No.	Project data	Data Description	Data type	Measurement Unit
		Data for the construction items group		
1	Natural Ground Preparations	Natural ground preparations price	Numerical	Iraqi dinar
2	Width	Road width	Numerical	Meter
3	Earthwork Layers	Earthwork embankment price	Numerical	Iraqi dinar
4	Width	Earthworks width	Numerical	Meter
5	Thickness	Earthworks thickness	Numerical	Meter
6	Granular Sub-Base Layer	Granular sub-base layer price	Numerical	Iraqi dinar
7	Width	Granular sub-base works width	Numerical	Meter
8	Thickness	Mixed gravel layer width	Numerical	Meter
9	Asphalt Concrete Base Layer	Asphalt concrete base layer price	Numerical	Iraqi dinar
10	Width	Paving width	Numerical	Meter
11	Thickness	Paving thickness	Numerical	Meter
12	Pipe Tunnel 60cm	Pipe tunnel installation works price	Numerical	Iraqi dinar
13	Granular Shoulder Layer	Granular shoulder works price	Numerical	Iraqi dinar
14	Width	Granular shoulder works width	Numerical	Meter
15	Thickness	Granular shoulder works thickness	Numerical	Meter
		Economic feature group		
1	GDP	Iraq's Gross Domestic Production per capita	Numerical	N/A
2	Unemployment Index	Iraq Unemployment Rate	Numerical	%
3	Inflation Index	Iraq Inflation Rate	Numerical	%
4	Oil Price	Crude oil price	Numerical	$
5	Dollar Exchange Rate	Dollar change	Numerical	$
6	Region	Location of the project 1. Khanaqin district 2. Al-Miqdadiya district 3. Baladruze district 4. Ba'quba district 5. Al Khalis district	Numerical	N/A
7	Year	Year of execution	Numerical	N/A

4. MODELING METHODOLOGY

The methods recommended for predicting the cost of rural road construction are described in this section. The methodology consists of four stages in general, as shown in Fig. 1. Stage (1) Data Collection, stage (2) Machine Learning (ML) Models, stage (3) Model Evaluation, and stage (4) Analysis Performance.

4.1 STAGE 1: DATA COLLECTION

This research collected the Bill of Quantities (BOQ) for about 3000 road construction projects in rural areas in the Diyala governorate from 2012 to 2021. These projects included many types of projects such as; new road construction, construction of asphalt pavement layers only, asphalt overlay, and pavement maintenance. After conducting a screening process for these projects, only those whose construction items were chosen were considered in this research while excluding all the others. This process is justified by the uniformity of construction items whenever they are adopted, which may facilitate the training process. The Department of Roads and Bridges in the Diyala government was the source, so the raw data has been obtained from it. However, these data were unsuitable for the proposed system to be trained on, so the researcher manually worked on it for three months and reshaped it for a CSV file to train the model.

BOQ: A schedule set by the employer's engineer according to the paragraphs must be implemented successively. For the state departments in Iraq, the process requires the formation of a committee to organize the inspection, which conducts the on-site inspection on the site, whether it is a construction or maintenance detection.

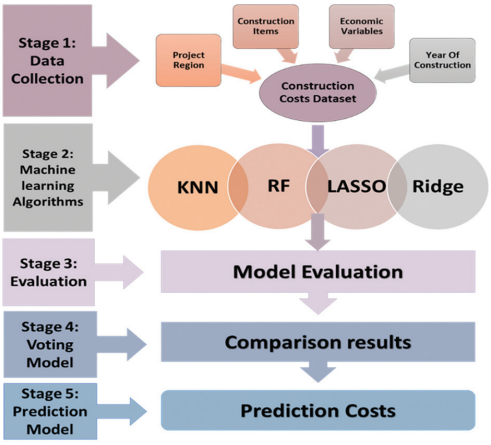


Fig. 1. Framework research cost roads regression.

STAGE 2: MACHINE LEARNING (ML) MODELS

ML approaches typically utilized are classification, clustering and regression [27]. The regression approach can be used to calculate the cost of construction. In general, any ML model consists of essentially three phases are explained as follows:

4.1.1 Preprocessing Phase

The primary goal of preprocessing is to turn raw data into a more suitable format for the predictor, allowing the prediction model to find patterns in the incoming data with ease.

I. Missing Data Preprocessing

Raw data was obtained from a legitimate government source and manually entered into an excel file, which was then prepared for use in the proposed project cost prediction model.

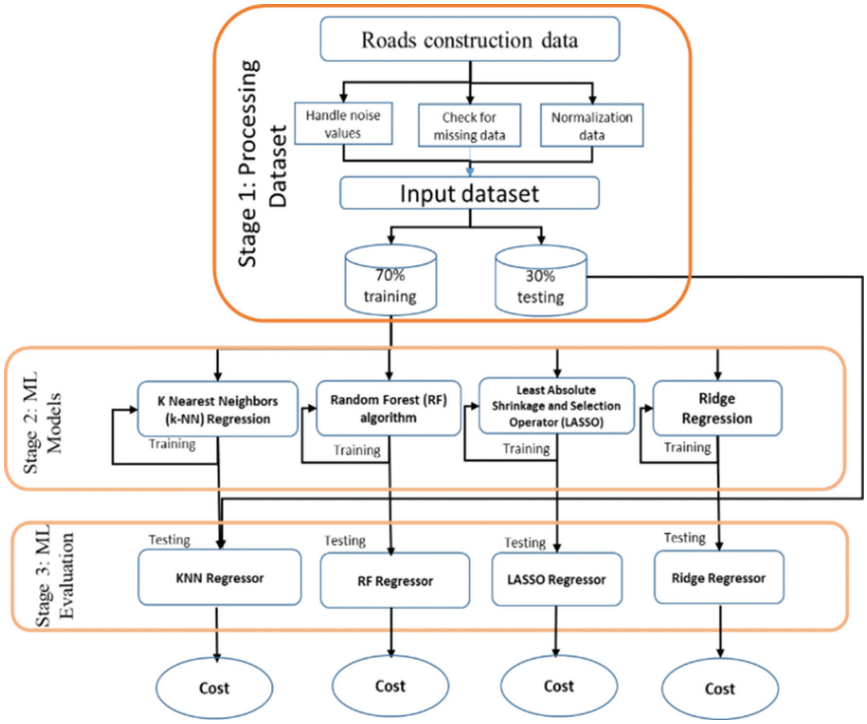


Fig. 2. Framework Machine Learning Models for cost estimation.

Therefore, there is no need for any compensatory approaches. After all, no compensating methods are required because there is no missing data following verification.

II. Data Normalization

Normalization aids in transforming an attribute's value into a limited set of values. It is the process of transmitting data to a specific range, such as 0 to 1 or -1 to 1. Although there are large differences in the values of different feature ranges, normalization is required. Data normalization, in which training time is started to access feature ranges of the same size, reduces training time.

4.1.2 Training Phase

In supervised learning, the classifier requires a training set that contains labeled samples of the domain with which it interacts to extract the requisite information and use it to predict future unlabeled inputs. Additionally, labeled data is necessary to test the classifier's performance by comparing the classifier's predictions to the actual classes of the inputs. The utilization of data was obtained from the training set for evaluation.

As indicated in Figure 2, 70% of the samples in the dataset were utilized for training, while the residual 30% were utilized for assessment in this research. Because the testing data is not included in the training data, this technique ensures unbiased evaluation by keeping a large ratio of data for evaluation purposes. The training starts by changing the parameters of each algorithm, as given in the next section, and then evaluating the outcome. For the remaining parameters, the operation is repeated.

4.1.3 Testing Phase

During the testing phase, four prominent supervised ML algorithms are utilized to determine effective and efficient forecast models for road-building costs. K Nearest Neighbors Regression (k-NN), Random Forest (RF), Ridge and Least Absolute Shrinkage and Selection Operator (LASSO) Regressions have been used. The turning and training parameters of each algorithm are explored to generate the finest feasible prediction outcomes, and many models of these algorithms are shown for (k-NN)it was test the efficiency for the most important parameter was the value of K as shown in table 2, and the performance for RF model was test for the n_estimators and max_depth parameters as in Table 3 illstraute.

Table 2. K-NN performance

Model setting	R^2	MAPE	RMSE	Time 1	Time 2
1	0.99	0.005	1494.9	0.01	0.01
2	0.99	0.005	1924.0	0.03	0.01
3	0.99	0.006	2190.3	0.03	0.01
4	0.99	0.007	2401.0	0.5	0.01
5	0.99	0.008	2740.3	0.01	0.01

Table 3. RF performance

Model setting	R^2	MAPE	RMSE	Time 1	Time 2
n_estimators =100 max_depth= None	0.99	0.005	3265.5	0.4	0.01
n_estimators =1 max_depth= 1	0.93	4.6	31798.0	0.04	0.00
n_estimators =1 max_depth= 6	0.95	4.33	28624.6	0.04	0.00
n_estimators =100 max_depth= 6	0.99	0.02	3620.57	0.2	0.01

4.2. STAGE 3: MODELS EVALUATION

1. Regression analysis is an important part of supervised ML since it involves predicting a continuous independent target from a set of predictor variables. Various studies employ the Mean Square Error (MSE) and its rooted variant (RMSE), including the Mean Absolute Error (MAE) and its percentage variant. However, these rates have one shortcoming: since their values might vary from zero to infinity, a single value does not tell anything about the regression's efficiency in regard to the ground truth distribution. Therefore, this research employed two rates that only produce a large score if most of the elements in a ground truth group are accurately predicted.[28].

2. R-squared (R^2) is the coefficient of determination described as the fraction of the dependent variable's variance that may be estimated by the independent variables. The degree to which the model fits the cost data is expressed as follows:[28][29].

$$\boldsymbol{R^2 = 1 - \left(\frac{SSE}{SST}\right)} \quad (5)$$

   in which SSE (sum of squares error) refers to the sum of squares of the residuals and SST (sum of squares total)refers to the total sum of squares. (Worst value = -∞; Best value = +1) [28].

3. MAPE is a regression model performance metric preferred for situations where relative variations are more highly relevant than absolute variations [30].

   Provided that x resembles the explanatory variables vector (the input to the regression model), y resembles the target variable as well as g denotes regression model, the MAPE of g is achieved by averaging the ratio over the data [31].

$$\boldsymbol{MAPE = \frac{|g(x)-y|}{|y|}} \quad (6)$$

4. ***MSE***: A risk metric related to the squared (quadratic) mistake or loss's predicted value [31]. If $\hat{y_i}$ denotes the predicted value of the i-$_{th}$ sample, as well as y_i denotes the corresponding true value, here, the MSE estimated over is presented as [31].

$$\boldsymbol{MSE\ (y, \hat{y}) = \frac{1}{n_{samples}} \sum_{i=0}^{n_{samples}-1} (y_i - \hat{y}_i)^2} \quad (7)$$

5. ***Root mean square error (RMSE)***: is the standard deviation of residuals (prediction errors). The RMSE is a measure of how spread out the residuals are, and the residuals are a measure of how distant the data points are from the regression line [31].

$$RMSE\ (y, \hat{y}) = \sqrt{\frac{1}{n_{samples}} \sum_{i=0}^{n_{samples}-1} (y_i - \hat{y}_i)^2}\ . \quad (8)$$

6. Training Time: The time it takes for a strategy to train the complete dataset and build the best-fit predictive model is referred to as (T1) [32].

$$T1 = end_{Time}^{Training} - Start_{Time}^{Training} \quad (9)$$

7. Testing time: (T2) is the time required for a technique to estimate the full dataset's construction costs [32].

$$T2 = end_{Time}^{Testing} - Start_{Time}^{Testing} \quad (10)$$

4.3 STAGE 4: ANALYSIS OF PERFORMANCE

R2, which evaluates how well the model matches the cost data, was used to assess model performance. As seen in table 1, R2 varies from 0 to 1. R2 values that are higher suggest better model performance. In part, the RMSE, which measures the average magnitude of an error, was employed to evaluate model performance. The RMSE should be as close to zero as feasible to indicate excellent model performance (for example, no error between real and anticipated costs). The performance of the model was assessed in part using MAPE, which determines if the model is more sensitive to relative than absolute fluctuations.

5. COMPARATIVE RESULT

For each of the four algorithms devised in this work, the extent of the modeling error was calculated, and comparison graphs of plots vs. output (for example, actual cost vs. predicted cost) were constructed. A diverse set of machine learning (ML) methods has been used to reach the optimal method for the road's construction real dataset proposed for the predictive cost model.

K Nearest Neighbors Regression (k-NN) was the first algorithm chosen to experiment with the instance-based algorithms. This method makes a decision using examples or instances of training data that the model considers necessary or essential. The R2 score value obtained by k-NN is 0.99, and the Mean Absolute Percentage Error (MAPE) is 0.006, but the Root Mean Square Error (RMSE) was very high 1485.6. Such algorithms frequently establish a database of example data and contrast incoming data to the database by utilizing a similarity measure to obtain the best match and produce a forecast. Random Forest (RF) ensemble techniques are models built of numerous weaker models that are individually trained and whose predictions are pooled in a particular manner to create the overall prediction. This data (RF) model failed to fit the data because of the linearity type of the data, so RMSE was 3181.8.

The Multicollinearity in this dataset leads to very high errors in the test phase in the K-NN and RF models. Multicollinearity is a condition to allow the correlation between the independent variables. Algorithms for regularization are another strategy (typically regression methods) that penalizes models for their complexity, preferring simpler models that are also stronger at generalizing. Ridge and Least Absolute Shrinkage and Selection Operator (LASSO) Regressions, the most popular algorithms for this method, have been utilized in this research. This kind of algorithm offered impressive results with the proposed dataset. The objective of lasso regression is to find the variables and regression coefficients that lead to a model with the least amount of prediction error. That is accomplished by imposing a constraint on the model parameters, which forces the sum of the absolute value of the regression coefficients to be smaller than a fixed value λ, hence shrinking the regression coefficients toward zero. The R2 score value obtained by LASSO Regression is 0.99, MAPE is 0.0002, and the RMSE is 0.09. Ridge Regression obtained the most elevated accuracy; in this model, it was tried to minimize the loss function, and the model was forced to find a balance between minimizing the residual sum of squares and minimizing the coefficients, which reached 0.000 for RMSE 1.00, R2, and the MAPE was 0.000. All models used the default parameters from the scikit-learn python library.

Time is the most important factor for calculating project costs, so this study focuses on the time factor for each ML model employed on the dataset, as shown in Table 4.

RF regression was the most terrible in training time than other models.

The time for training and testing is gradually reduced. The more accurate the model, the less error it is. That is evident in all other models, and when the model is perfect and gives a zero error, the time is ideal for training and testing, which applies to our proposed model, the Ridge Regression.

Table 4. Models performance

Regressors	T1 (s)	T2 (s)	R^2	MAPE	MSE	RMSE
Random Forest	0.53554	0.0 0978	0.99	0.0 07	10123 893.2	31 81.8
K-NN	0.03398	0.0 1499	0.99	0.0 06	22070 83.8	14 85.6
Lasso	0.31494	0.0 0000	0.99	0.0 0002	0.0 0383	0.0 6190
Ridge	0.40817	0.0 0000	1.00	0.0 00	0.0 000	0.0 0000

The following Figures show errors or differences between the predicted labels and the actual labels for road construction costs based on machine learning regressions. The RF modulator and the K-NN slope have errors in the prediction cost that do not lie on a

straight regression line. However, the other points' cost estimates lie correctly on the same line regression, as shown in Figures 3 and 5. LASSO and Ridge have presented the errors and actual road cost estimation in Figures 4 and 6. These regressions are shown as the accurate and great model for predicting costs for road construction based on error, and actual points lie precisely on one-line regression.

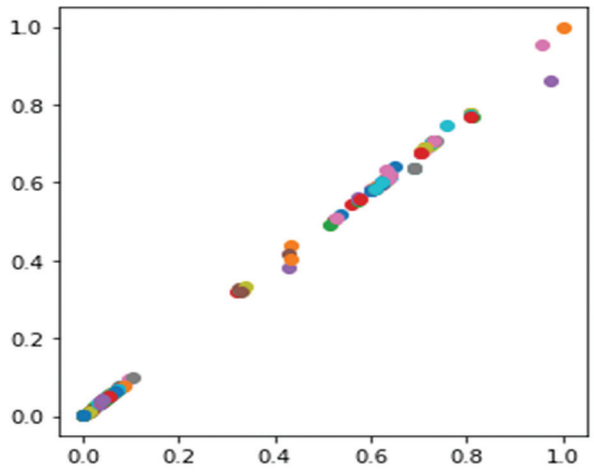


Fig. 3. RF Regressor

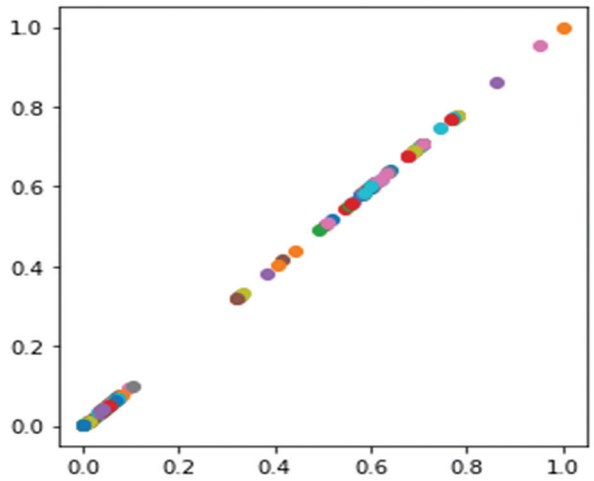


Fig. 4. LASSO Regressor

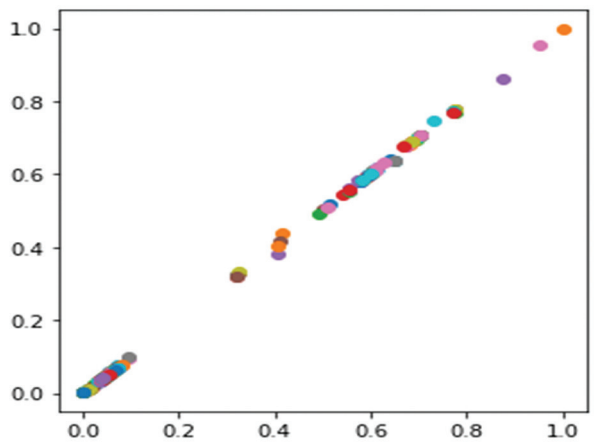


Fig. 5. K NN Regressor.

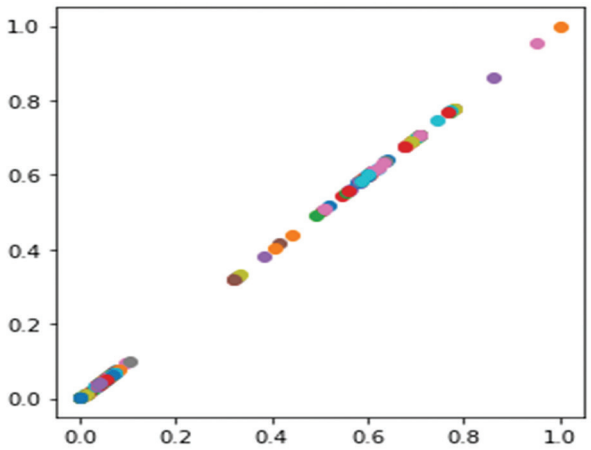


Fig. 6. Ridge Regressor.

6. CONCLUSIONS

Cost prediction is necessary for civil engineers to need ample time and action. This study collected a real road dataset for ten years. Furthermore, it examined the performance of several machine learning (ML) modeling for the prediction of road projects. The research aimed to advance machine learning techniques in estimating road construction. The suggested computing method is all-encompassing and can be applied to various construction projects in any city or region. The model has been presented using information gathered from a town. The proposed algorithm can be examined for additional building projects and areas.

On the road cost construction testing, the evaluation scale the Ridge model received the most significant estimate among the four models. In contrast, the other ML models had an acceptable level of cost estimation, except for the K Nearest Neighbors Regression (k-NN) model and the Random Forest (RF) model, which showed high errors in the cost prediction because of the linearity type for this data. This research directs a new direction for the ML approach to provide an economically reasonable cost with minimal effort to manage projects by artificial intelligence (AI) to achieve important goals for a road business. Additional research in the future might expand this analysis to include the use of deep learning algorithms and to collect data for different fields in construction engineering, another source to Increase the dataset to generate more precise cost estimation models for various objectives.

7. REFERENCES

[1] S. Petrusheva, D. Car-Pušić, V. Zileska-Pancovska, "Support Vector Machine Based Hybrid Model for Prediction of Road Structures Construction Costs", Proceedings of the IOP Conference Series: Earth and Environmental Science, Vol. 222, No. 1, 2019, pp. 1-11.

[2] I. Karaca, D. D. Gransberg, H. D. Jeong, "Improving the Accuracy of Early Cost Estimates on Transportation Infrastructure Projects", Journal of Management in Engineering, Vol. 36, No. 5, 2020, p. 04020063.

[3] M. Baek, B. Ashuri, "Spatial regression analysis for modeling the spatial variation in highway construction costs", Resilient Structures and Sustainable Construction, Georgia Institute of Technology, Atlanta, 2017.

[4] A. J. Ogungbile, A. E. Oke, K. Rasak, "Developing cost model for preliminary estimate of road projects in Nigeria", International Journal of Sustainable Real Estate and Construction Economics, Vol. 1, No. 2, 2018, pp. 182-199.

[5] A. Jaafari, I. Pazhouhan, P. Bettinger, "Machine learning modeling of forest road construction costs", Forests, Vol. 12, No. 9, 2021, p. 1169.

[6] L. B. Barros, M. Marcy, M. T. M. Carvalho, "Construction Cost Estimation of Brazilian Highways Using Artificial Neural Networks", International Journal of Civil Engineering, Vol. 7, No. 3, 2018, pp. 283-289.

[7] K. Tijanić, D. Car-Pušić, M. Šperac, "Cost estimation in road construction using artificial neural network", Neural Computing and Applications, Vol. 32, No. 13. 2020, pp. 9343-9355,

[8] G. Mahalakshmi, C. Rajasekaran, "Early cost estimation of highway projects in India using artificial neural network", Sustainable construction and building materials, Springer, Singapore, 2019, pp. 659-672.

[9] F. M. S. AL-Zwainy, I. A.-A. Aidan, "Forecasting the Cost of Structure of Infrastructure Projects Utilizing Artificial Neural Network Model (Highway Projects as Case Study)", Indian Journal of Science and Technology, Vol. 10, No. 20, 2017, pp. 1-12.

[10] I. Peško et al. "Estimation of costs and durations of construction of urban roads using ANN and SVM", Complexity, Vol. 2017, 2017, pp. 1-13.

[11] N. Suneja, J. P. Shah, Z. H. Shah, M. S. Holia, "A neural network approach to design reality oriented cost estimate model for infrastructure projects", Reliability: Theory & Applications, Vol. 16, No. 1, 2021, pp. 254-263.

[12] M. Juszczyk, "On the search of models for early cost estimates of bridges: An SVM-based approach", Buildings, Vol. 10, No. 1. 2020, pp. 1-17.

[13] I. Mahamid, "Early cost estimating for road construction projects using multiple regression techniques", Australasian Journal of Construction Economics and Building, Vol. 11, No. 4, 2011, pp. 87-101.

[14] T. C. D. Lucas, "A translucent box: interpretable machine learning in ecology", Ecological Monographs, Vol. 90, No. 4, 2020, pp. 1-55.

[15] Y. Cao, B. Ashuri, M. Baek, "Prediction of Unit Price Bids of Resurfacing Highway Projects through Ensemble Machine Learning", Journal of Computing in Civil Engineering, Vol. 32, No. 5, 2018, p. 04018043.

[16] M. Flah, I. Nunez, W. Ben Chaabene, M. L. Nehdi, "Machine Learning Algorithms in Civil Structural Health Monitoring: A Systematic Review", Archives of Computational Methods in Engineering, Vol. 28, No. 4. 2021, pp. 2621-2643.

[17] I. Triguero, D. García-Gil, J. Maillo, J. Luengo, S. García, F. Herrera, "Transforming big data into smart data: An insight on the use of the k-nearest neighbors algorithm to obtain quality data", Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, Vol. 9, No. 2, 2019, pp. 1-24.

[18] F. Martínez, M. P. Frías, F. Charte, A. J. Rivera, "Time Series Forecasting with KNN in R : the tsfknn Package", The R Journal, Vol. 11, No. 2, 2019, pp. 229-242.

[19] G. C. McDonald, "Ridge regression", Wiley Interdisciplinary Reviews: Computational Statistics, Vol. 1, No. 1, 2009, pp. 93-100.

[20] C. Saunders, A. Gammerman, V. Vovk, "Ridge Regression Learning Algorithm in Dual Variables", Proceedings of the Fifteenth International Conference on Machine Learning, 1998, pp. 515-521.

[21] A. Afzal, S. Alshahrani, A. Alrobaian, A. Buradi, S. A. Khan, "Power plant energy predictions based on thermal factors using ridge and support vector regressor algorithms", Energies, Vol. 14, No. 21, 2021, pp. 1-22.

[22] L. E. Melkumova, S. Y. Shatskikh, "Comparing Ridge and LASSO estimators for data analysis", Procedia Engineering, Vol. 201, 2017, pp. 746-755.

[23] J. F. H. Tibshirani, "Regularization Paths for Generalized Linear Models via Coordinate Descent", Journal of Statistical Software, Vol. 33, No. 1, 2010, pp. 1-22.

[24] S. J. Kim, K. Koh, M. Lustig, S. Boyd, D. Gorinevsky, "An interior-point method for large-scale ℓ1-regularized least squares", IEEE Journal of Selected Topics in Signal Processing, Vol. 1, No. 4, 2007, pp. 606-617.

[25] F. Hidayat, T. M. S. Astsauri, "Applied random forest for parameter sensitivity of low salinity water Injection (LSWI) implementation on carbonate reservoir", Alexandria Engineering Journal, Vol. 61, No. 3, 2022, pp. 2408-2417.

[26] H. Tong, B. Liu, S. Wang, "Software defect prediction using stacked denoising autoencoders and two-stage ensemble learning", Information and Software Technology, Vol. 96, 2018, pp. 94-111,

[27] S. B. Jha, R. F. Babiceanu, V. Pandey, R. K. Jha, "Housing Market Prediction Problem using Different Machine Learning Algorithms: A Case Study", arxiv.org/abs/2006.10092, 2020.

[28] D. Chicco, M. J. Warrens, G. Jurman, "The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation", PeerJ Computer Science, Vol. 7, 2021, pp. 1-24.

[29] H. H. Elmousalami, "Artificial Intelligence and Parametric Construction Cost Estimate Modeling: State-of-the-Art Review", Journal of Construction Engineering and Management, Vol. 146, No. 1, 2020, p. 03119008.

[30] H. H. Elmousalami, "Comparison of Artificial Intelligence Techniques for Project Conceptual Cost Prediction: A Case Study and Comparative Analysis", IEEE Transactions on Engineering Management, Vol. 68, No. 1. 2021, pp. 183-196.

[31] A. de Myttenaere, B. Golden, B. Le Grand, F. Rossi, "Mean Absolute Percentage Error for regression models", Neurocomputing, Vol. 192, 2016, pp. 38-48.

[32] Z. K. Maseer, R. Yusof, N. Bahaman, S. A. Mostafa, C. F. M. Foozy, "Benchmarking of Machine Learning for Anomaly Based Intrusion Detection Systems in the CICIDS2017 Dataset", IEEE Access, Vol. 9, 2021, pp. 22351-22370.

Driving pattern analysis to determine driver behaviours for a local authority based on the cloud using OBD II

Original Scientific Paper

Siddhanta Kumar Singh
Department of Computer Science and Engineering
Mody University
Lakshmangarh, India
singhsiddhant@yahoo.com

Ajay Kumar Singh
Department of Computer Science and Engineering
Mody University
Lakshmangarh, India
ajay.kr.singh07@gmail.com

Abstract *– Aggressive driving is the leading cause of road accidents, which mainly results from driving behaviour that endangers drivers and the people around them. Such driver behaviours should be identified by local authorities to correct the behaviours or understanding the root cause of accidents. Data recorded using the On Board Diagnostic (OBD) II device can be analyzed to determine such behaviours. A sudden change in the manoeuvring of a vehicle indicates aggressive driving behaviour, which eventually yields non-uniform parameter values returned by the engine control unit (ECU) system without any specific reason. In this study, real-time data were recorded from ECU by using OBD II and an accelerometer. Artificial Intelligence was used to group different types of data to identify behaviours based on the similarity of data points. This study aimed to identify such drivers and reduce the risk of accidents. Driving behaviours were categorized as bad, normal, and aggressive. Because clustering is based on crowded data, which signifies similar driving patterns most of the time in the course of recording, this study used density-based spatial clustering of applications with noise unsupervised learning algorithm. The system sends data to the cloud, enabling the authorities to access it from any place for further action. ANOVA was conducted using IBM's Statistical Package for Social Sciences to compare and determine the most favourable method to collect data.*

Keywords*: Statistical Package for the Social Sciences, Engine Control Module, On Board Diagnostic II, Electronic Control Unit, Controller Area Network*

1. INTRODUCTION

Vehicles are equipped with many sensors that provide beneficial real-time information, such as that on speed, rpm, acceleration, and fuel consumption. Moreover, these data reflect the vehicle's condition and driver's behaviour. Previously, cars used On Board Diagnostic (OBD) I instead of OBD II protocols. OBD I protocols are manufacturer specific. The port and scanner differ for varying companies. This inconsistency led to the development of generic OBD II protocols. This study identifies the driving behaviour based on the driving pattern, considering engine OBD II parameters. Moreover, this study analyses the driving pattern based on OBD II data, such as speed and rpm. The findings of this study can help determine the previous driving behaviour and accordingly design measures based on travel history.

The direct association of driving behaviour with the driving style should be examined because it affects a vehicle's fuel consumption, emission, and safety. The OBD [1] helps to collect real-time data for post-analysis. The California Air Resources Board and the United States Environmental Protection Agency developed the OBD. The Society for Automotive Engineers and International Organization for Standardization standardized the OBD in the United States and worldwide. The iSaddle OBD-II scanner as shown in Fig. 1 has a small blue tooth adapter that supports all OBD II protocols [2]. We can determine a vehicle's protocol by examining the data link connector pinout as shown in Fig. 2. The CAB bus uses a differential signal system instead of a binary system as shown in Fig. 3.

The physical layer of the CAN Bus protocol uses differential transmission on a twisted pair medium. The messages are of eight data bytes and are protected by a checksum. No exact address is provided in messages; each message has a numeric value that controls its priority on the bus and may identify its contents. The lower the numerical ID in the message format is, the higher the message priority.

Secure communication channels should be ensured for data transfer. Databases placed on file servers should be protected, and users that possess and use those data should be controlled. Technical security solutions for physical and software protection with security procedures for redundancy and backup automation should be of high quality. However, the ongoing cyber war is still not suppressed. A possible reason for this is technical security solutions because they rarely consider the effect of the human factor on the security level of the system.

The human factor is the weakest element in the security chain because the internal threat is among the top information security problems [1].

Empirical studies examining the amount of human influence in the field of IT security are lacking [18]. Some existing empirical studies have analyzed user perception, behaviour, and attitude towards computer ethics and information security [3-5] because computer security and ethics are the essential components of a management information system [6].

Fig. 1. iSaddle OBD II Scanner

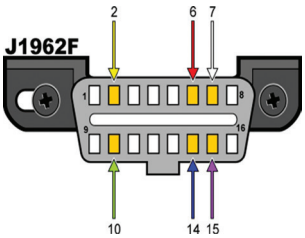


Fig. 2. Pinout diagram of OBD II port

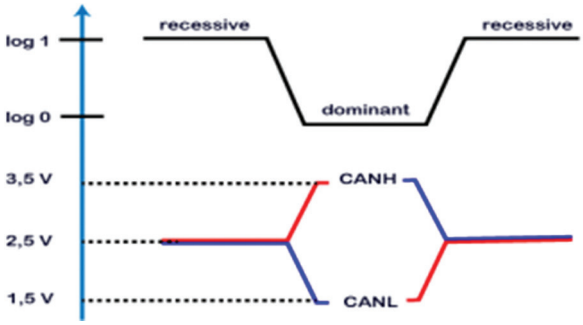


Fig. 3. Dominant and recessive CAN Bus signals

Bit-wise arbitration of the CAN-ID can resolve network access conflicts. All connected nodes observe the bus level bit-for-bit. All nodes transmit a recessive level and detect a dominant level, lose bus arbitration, and transit into the listening mode.

Fig. 4 presents how the ECU continuously broadcasts different sensor signals on the CAN bus identified by CAN ID (Standard ID is a hexadecimal value) that addresses the data and priority of arbitration. The lower the value is, the higher the priority. The data length is provided in bytes, and the data are also provided in bytes. The raw data are extracted by coupling the microcontroller with MCP2515 IC and the vehicle's OBD II port. Fig. 5 presents the equipment setup and connection.

```
13:46:53.432 -> Entering Configuration Mode Successful!
13:46:54.057 -> Setting Baudrate Successful!
13:46:54.057 -> MCP2515 Initialized Successfully!
13:46:54.057 -> MCP2515 Library Receive Example...
13:46:54.057 -> Standard ID: 0x158       DLC: 8  Data: 0x00 0x000x000x000x000x000x00 0x0A
13:46:54.057 -> Standard ID: 0x17C       DLC: 8  Data: 0x00 0x000x000x00 0x10 0x00 0x00 0x03
13:46:54.057 -> Standard ID: 0x158       DLC: 8  Data: 0x00 0x000x000x000x000x000x00 0x19
13:46:54.057 -> Standard ID: 0x1AB       DLC: 3  Data: 0x00 0x00 0x5D
13:46:54.057 -> Standard ID: 0x294       DLC: 8  Data: 0x00 0x000x000x00 0x48 0x89 0x00 0x1B
```

Fig. 4. Sample CAN raw data

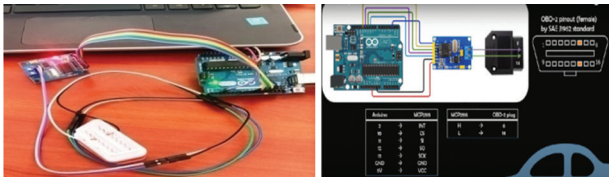


Fig. 5. Arduino, MCP, and OBD II setup

2. RELATED WORKS

Kawtar et al. described the vehicle-, management-, and driver-oriented classes of application. They mainly focused on the driver-oriented class [3], which comprises accident prevention, driving style assessment, and driver intent prediction. They used descriptive statistics and Bayesian classifiers to present the results. Peppeset al. [4] described a platform that combines machine and deep learning algorithms utilizing clustering techniques together with open-source-based tools to gather, store, process, analyze, and correlate different data retrieved from vehicles.

Navneeth et al. [5] examined the driver's profile and behaviour. Ameen et al. [6] identified driving behaviours to reduce the risk of accidents based on real-time data recorded from vehicles and reference data provided by previous researchers. The t-test was used to compare mean values between groups, and statistical analyses were performed using Statistical Package for the Social Sciences (SPSS).

Shaikh et al. [7] developed an Android application to alert any abnormality or anomaly in driving behaviour by using OBD II with Bluetooth and a Wifi connection. Hermawan et al. [8] examined driving behaviours and various methods to obtain OBD-II data for analyzing, modelling, and evaluating systems. Pan et al. [9] used the logistic regression model and examined behavioural parameters that affect a vehicle's risk situation and factors that affect safe driving. Ameen et

al. [10] determined driving data, such as speed, rpm, and throttle position, by using OBD II. They proposed a method to develop a driving behaviour classification by using severity stratification of the acceleration model to determine safe driving. They mentioned an acceleration level of approximately ±2 m/s^2 for safe driving behaviour and ±4 m/s^2 for aggressive drivers with a risk of collision. Uvarov et al. [11] identified drivers with sensors listed in SAE J1979 specification, but the accuracy declined by approximately 15%. Sivaraj et al. [12] compared the standard limits of different OBD II parameters, such as speed, acceleration, retardation, and jerk. The telegram application they developed notified of any deviation observed in the driving behaviour. Jiang et al. [13] used a GPS-enabled smartphone app with a zero-inflated negative binomial regression model to investigate drivers' overspeeding violation behaviour for safety diagnosis and traffic warnings. Zhang et al. [14] analyzed the driving behaviour by using sparse automatic encoders and explored data to detect abnormal and aggressive behaviour. Regression analysis was performed to investigate the relationship between aggressive driving and road facilities.

Xiang et al. [15] proposed a new hybrid model consisting of cloud and the Elman neural network (CM-ENN) for predicting dangerous driving behaviour based on vehicle motion state estimation and passengers' subjective feeling scores. Dixit et al. [16] developed an OBD II vehicular data acquirement and analytics system for trip analysis and vehicle diagnostics and to monitor real-time driving behaviour based on cloud data. Jeon et al. [17] used LoRa communication to transmit vehicle operation information periodically to the server for analysis.

Meseguer et al [18] designed a real-time monitoring application to examine the correlation between the driver's physiological and the vehicle's diagnostic data. He et al. [19] mounted OBD II devices in HDVs in China to gather high-precision and sampling frequency data for verification. A driving behaviour portrait approach was proposed based on the driving behaviour and frequency and ranking of drivers' typical driving patterns. Agrawal et al. [20] proposed an analytical solution by using cluster analysis to detect safe or rash driving behaviour. Sonawane et al. [21] developed a framework that can be used for the clustered analysis of driving patterns by using data acquired from the OBD-II device. Massoud et al. [22] developed serious games (SGs) for obtaining information on drivers' trips by using a human driving profiling algorithm. Their developed system helped the driver in determining fuel efficiency. Trindade et al. [23] gathered data from vehicles' and smartphones' sensors. Their application enabled the evaluation of driving behaviour.

Fu et al. [24] developed the driving behaviour risk prediction neural network for prediction based on distracted driving behavioural data.

Mohammed et al. [25] compared driving behaviour between day and night, between weekends and weekdays, and among different road types (east, north, and south). They used the digital dashboard GPS pro app to collect data. The independent t-test and one-way ANOVA were used to analyze data. Alluhaibi et al. [26] proposed several methods to detect driving behaviours and identified each method's advantages and disadvantages. Wang et al. [27] used machine learning techniques to investigate drivers' behaviours in different traffic scenarios.

3. METHODOLOGY AND RESULTS

The app stored data in the Firebase cloud in the csv format. A part of the sample data is presented in Table 1.

Table. 1. Driving dataset

Device Time (sec)	G(x)	G(y)	G(z)	Speed (GPS) (km/h)	Speed (OBD) (km/h)	RPM
0	-3.3	3.59	6.02	12.58	10	783
1	1.35	2.67	9.05	11.79	10	783
2	1.17	2.49	9.57	11.11	10	784
3	1.45	2.27	9.76	10.62	9	812
4	1.34	1.88	9.57	10.17	9	1437.5

ANOVA was used to determine the most favourable approach to gathering driving data [28]. IBM's SPSS was used to compare means between groups for the same continuous dependent variable. The independent sample t-test was conducted to compare the sample mean values between the two groups. In the case of more than two groups, we used one-way ANOVA because it has one independent variable for which three sample data at three times points for three vehicles can be gathered. The sample groups were labelled 1, 2, and 3, respectively, and the dependent variable was speed. Table 2 presents the means, standard deviations, and standard errors of different sample data computed using SPSS. No significant differences were observed in data collected from various sources. The findings of the test of homogeneity of variance for different samples are presented in Table 3.

Table 2. Description of car speeds

Descriptives								
Speed								
					95% Confidence Interval for Mean			
	N	Mean	Std. Deviation	Std. Error	Lower Bound	Upper Bound	Min	Max
1	152	.4290	.20266	.01644	.3966	.4615	.21	1.00
2	209	.4425	.24985	.01728	.4084	.4765	.00	.97
3	108	.4823	.16368	.01575	.4510	.5135	.15	.79
Total	46	.4473	.21799	.01007	.4275	.4671	.00	1.00

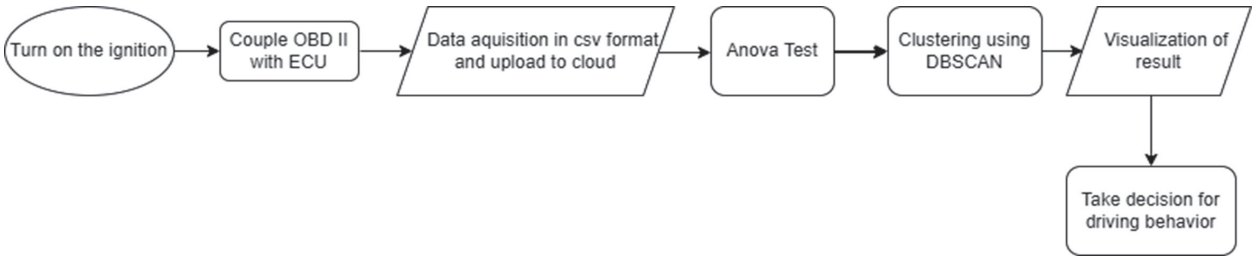


Fig. 6. The workflow design

One-way ANOVA was performed to compare speeds among groups 1, 2, and 3.

Table 3. Homogeneity of variances

Tests of Homogeneity of Variances					
		Levene Statistic	df1	df2	Sig.
Speed	Based on Mean	22.540	2	466	<.001
	Based on Median	20.792	2	466	<.001
	Based on the Median and with adjusted df	20.792	2	401.428	<.001
	Based on trimmed mean	22.841	2	466	<.001

The test was conducted by comparing the variation between the sample means and variation within each sample by using the following formulations. The difference in the group means was tested by partitioning the total variation into two components.

Fig. 6 presents the workflow design. The design methodology involved data acquisition from the ECU y using an OBD II scanner connected to a smartphone with Bluetooth and uploading the data to Firebase cloud storage. ANOVA was performed to examine captured data, and an unsupervised learning algorithm was employed to visualise behavioural data.

Variation of group means from the overall mean was calculated asthe $\bar{y}_{.j}$- $\bar{y}_{..}$(variation between groups), where $\bar{y}_{.j}$ is the sample mean of group j and $\bar{y}_{..}$ is the overall sample mean. Variation of observations in each group from their group mean estimates, y_{ij}- $\bar{y}_{.j}$ (variation within group).

$$\sum_i \sum_j \left(y_{ij} - \bar{y}_{..}\right)^2 \quad (1)$$

$$\sum_{j=1}^{k} n_j \left(\bar{y}_{\cdot j} - \bar{y}_{..}\right)^2 \quad (2)$$

$$\sum_i \sum_j \left(y_{ij} - \bar{y}_{.j}\right)^2 \quad (3)$$

ANOVA partitions the total sum of squares (SST) in eq. 1 into the sum of squares due to the between-group effect (SSR), as shown in eq. 2, and sum of squared errors (SSE), as in eq. 3.

$$\sum_i \sum_j \left(y_{ij} - \bar{y}_{..}\right)^2 = \sum_{j=1}^{k} n_j \left(\bar{y}_{\cdot j} - \bar{y}_{..}\right)^2 + \sum_i \sum_j \left(y_{ij} - \bar{y}_{.j}\right)^2 \quad (4)$$

where

$$SSR = \sum_{j=1}^{k} n_j \left(\bar{y}_{\cdot j} - \bar{y}_{..}\right)^2 \quad (5)$$

$$SSE = \sum_i \sum_j \left(y_{ij} - \bar{y}_{.j}\right)^2 \quad (6)$$

$$SST = \sum_i \sum_j \left(y_{ij} - \bar{y}_{..}\right)^2 \quad (7)$$

Where,

n_j is the size of the sample.

k is sample no

$df_{between}$(Degree of freedom) = k-1

df_{within}(Degree of freedom) = n-k

$MS_{between}$(Means Square) = SSR/$df_{between}$

MS_{within}(Means Square) = SSE/df_{within}

F(Anova coefficient) = $MS_{between}/MS_{within}$

Sig. was calculated from the F distribution table for the two degrees of freedom and F value.

No significant differences in Sig.'s values were greater than 0.05 as given in Table 4.

Table 4. ANOVA Test

ANOVA					
Speed					
	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	.188	2	.094	1.982	.139
Within Groups	22.052	466	.047		
Total	**22.240**	**468**			

The post hoc tests indicate the values of Sig. after comparing with the values of one group with the other two groups as shown in Table 5; these values differed significantly.

Three driving experiments were conducted in the city area by using three vehicles. The scatter plots are shown in Fig. 7(a), 7(b), and 7(c), and tests were conducted to measure the consistency of the data collected from the cloud. In this experiment, Honda Brio, Hyundai Accent, and Tata Tiago were tested under different traffic conditions.

We observed some pattern similarities and deduce behavioural outcomes from these patterns.

Table 5. Post Hoc Tests

Post Hoc Tests / Multiple Comparisons						
Dependent Variable: Speed						
LSD(Least Significant Difference)						
					95% Confidence Interval	
(I) Cars	(J) Cars	Mean Difference (I-J)	Std. Error	Sig.	Lower Bound	Upper Bound
1	2	-.01341	.02319	.563	-.0590	.0322
	3	-.05321	.02738	.053	-.1070	.0006
2	1	.01341	.02319	.563	-.0322	.0590
	3	-.03980	.02578	.123	-.0905	.0109
3	1	.05321	.02738	.053	-.0006	.1070
	2	.03980	.02578	.123	-.0109	.0905

In the three figures, speed is plotted on the x-axis and rpm on the y-axis. A major concentration of points was noted in the three groups except for some outliers, which eventually led to three patterns of driving.

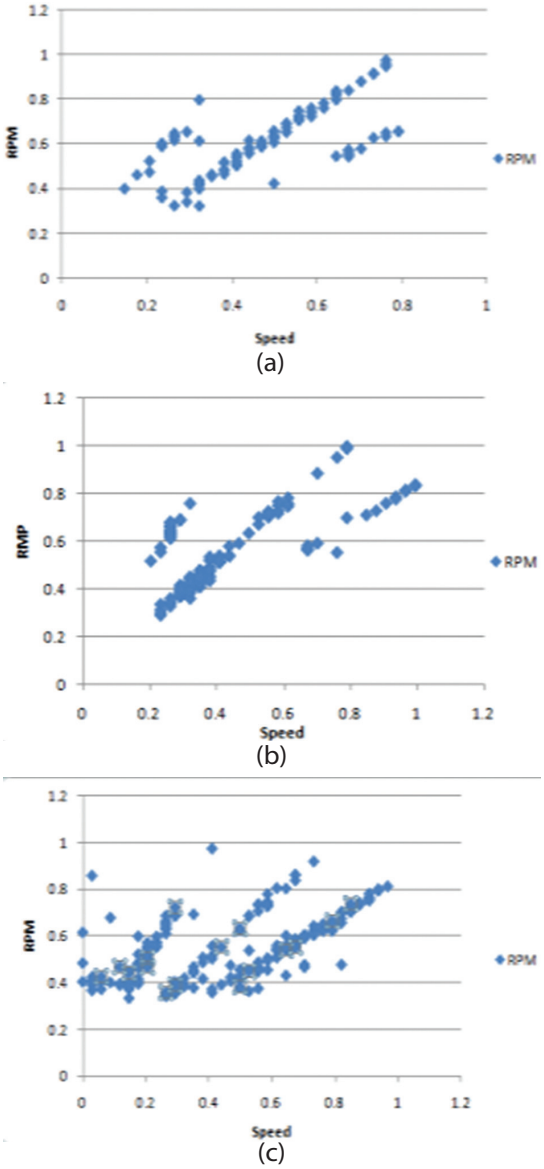


Fig. 7. Scatter plot of sample data

The dataset underwent a MinMaxScaler operation of Python. It returned a default value from 0 to 1 for both axes and preserved the original shape of the distribution of the dataset used. This process did not modify any information embedded in the original dataset. Moreover, the MinMaxScaler did not reduce the importance of outliers and noise.

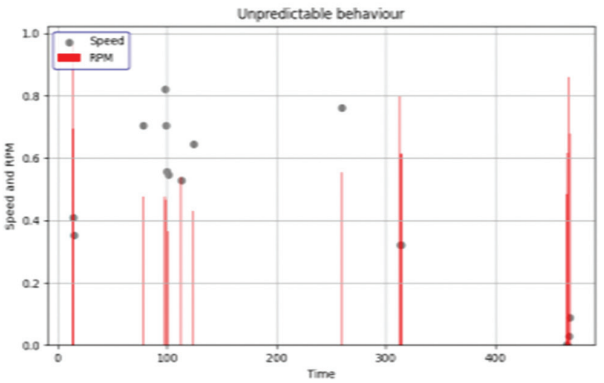
```
Total clusters in DBSCAN are= [-1  0  1  2]
Silhouette Score with 4 clusters is 0.41143794599627603
Intercept= [0.71741061] Coefficient= [[-0.30420312]]
Intercept= [0.27213606] Coefficient= [[1.32800002]]
Intercept= [0.01722772] Coefficient= [[1.24074742]]
Intercept= [0.02389404] Coefficient= [[0.8126347]]
```

Fig. 8. Clustering and regression parameters

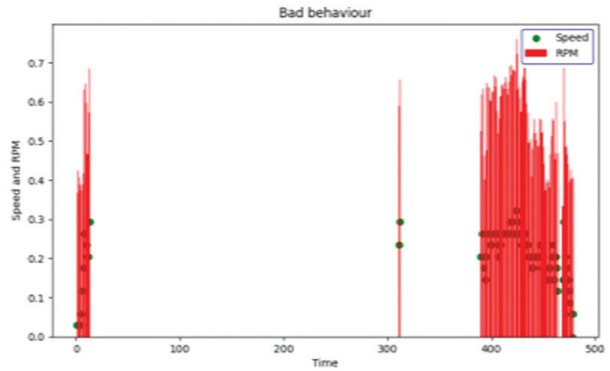
A linear regression line was fit using Python machine learning for different clusters given by various data frames, and we examined the nature of these regression lines. The coefficients and intercepts of these lines are presented in Fig. 8.

The −1 element in the list of the clustered output denotes outliers whose values are unpredictable, and regression coefficients had a negative value and a higher intercept. The remaining intercepts were small near zero when x was equal to 0. The positive coefficient value indicated an increase in rpm caused an increase in speed. The clusters were composed of different driving patterns, which depicted other driving behaviours, and regression lines were fitted with clustered points for each pattern. As presented in Fig. 9(a), 9(b), 9(c), and 9(d), driving patterns were classified as unpredictable, bad, normal, and good, respectively. The silhouette score was computed to test the accuracy of clusters. Normal, good, bad, and unpredictable driving was based on the speed and rpm of OBD II data recorded and eventually affected by comparative fuel consumption, which depends on throttling.

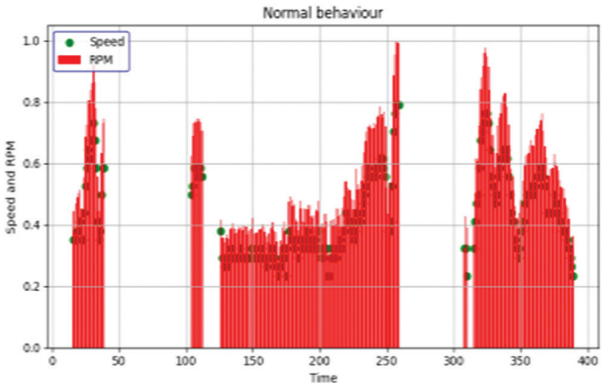
As presented in Fig. 9(e), the green marker represents speed and the red marker shows rpm. Driving behaviours were visualised on the basis of the relative positions of one marker above or below the other, as indicated in the algorithm.



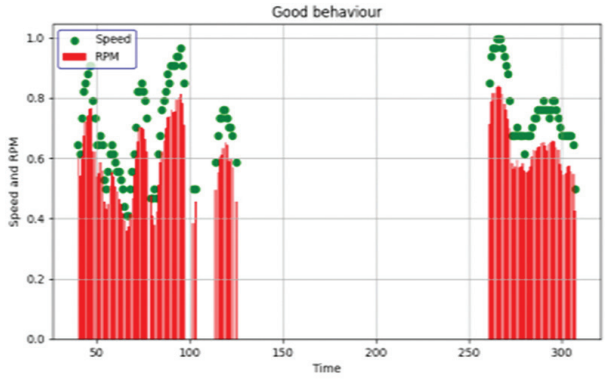
(a)



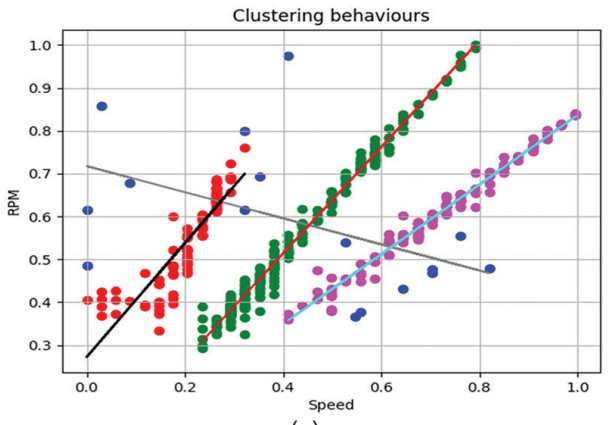
(b)



(c)



(d)



(e)

Fig. 9. Visualisation of driving behaviours

From the workflow design mentioned in the methodology section, driving behaviours can be computed as a function f(x,y) at time t on x and y for speed and rpm as

$f(x_t, y_t)$ = Unpredictable , if $x_t \not\approx y_t$

Bad, if $x_t << y_t$

Normal, if $x_t \approx y_t$

Good, if $x_t >> y_t$

We determined different driving patterns mathematically by using the aforementioned function and graph presented in Fig. 9(e).

4. CONCLUSION

Real-time OBD II data were gathered over a long period and sent to the server. Some of these analyses were performed within the app. Complex analyses were performed at the backend, that is, the cloud server. Data were forwarded to Firebase cloud storage. Once gathered data were received by the local authority or owner of the vehicle, they could determine the driver's behaviour for further action or reform the driver. The analysis can be conducted at the app level without the need to send the large data to the server. The analysis could be performed using the app, and only the resultant behavioural information was forwarded to the concerned party. This process can reduce the large size of OBD II data sent to the cloud.

5. REFERENCES:

[1] M. M. O. Abolarin, M. Sunday, "Specifications and Analysis of Digitized Diagnostics of Automobiles: A Case Study of on Board Diagnostic (OBD II)", International Journal of Engineering Research & Technology, Vol. 9 No. 1, 2020.

[2] J. V. Moniaga, S. R. Manalu, D. A. Hadipurnawan, F. Sahidi, "Diagnostics vehicle's condition using obd-ii and raspberry pi technology: study literature", Proceedings of the 2nd International Conference on Computing and Applied Informatics, 2017.

[3] K. Zinebi, N. Souissi, K. Tikito, "Driver Behaviour Analysis Methods: Applications oriented study", Proceedings of the 3rd International Conference on Big Data, Cloud and Applications, 2018.

[4] N. Peppes, T. Alexakis, E. Adamopoulou, K. Demestichas, "Driving Behaviour Analysis Using Machine and Deep Learning Methods for Continuous Streams of Vehicular Data", Sensors, Vol. 21, No. 14, 2021.

[5] S. Navneeth, K. Prithvil, N. Sri Hari, R. Thushar, M. Rajeswari, "On-Board Diagnostics and Driver Profiling", International Research Journal of Engineering and Technology, Vol. 7, No. 6, 2020.

[6] H. A. Ameen, "Driving Behaviour Identification based on OBD Speed and GPS Data Analysis", Advances in Science Technology and Engineering Systems Journal, Vol. 6, No. 1, 2021, pp. 550-569.

[7] M. K. Shaikh, S. Palaniappan, F. Ali, M. Khurram, "Identifying driver behaviour through OBD II using android application", PalArch's Journal of Archaeology of Egypt/Egyptology, Vol. 17, No. 7, 2020.

[8] G. Hermawan, E. Husni, "Acquisition, Modeling, and Evaluating Method of Driving Behaviour Based on OBD-II: A Literature Survey", IOP Conference Series: Materials Science and Engineering, 2020.

[9] Y. Pan, T. Yu, R.-S. Cheng, "Using OBD-II data to explore driving behaviour model", Proceedings of the International Conference on Applied System Innovation, 2017, pp. 1816-1818.

[10] H. A. Ameen et al. "Identification of Driving Safety Profiles in Vehicle to Vehicle Communication System Based on Vehicle OBD Information", Recent Advances in IoT and Cyber/Physical Security, Vol. 2, No. 5, 2021.

[11] K. Uvarov, A. Ponomarev, "Driver Identification with OBD-II Public Data", Proceedings of the 28th International Conference of Fruct Association, 2021, pp. 495-501.

[12] D. Sivaraj, S. H. Kumar, D. Jogarao, S. Dutta, K Ezhumalai, E. Rabjor, "Analysis of Driving Behaviour for Improved Safety in Commercial Fleet Management using Onboard Diagnostics (OBD-II)", Proceedings of the Smart Technologies, Communication and Robotics, 2021, pp. 1-5.

[13] Y. Jiang, J. Zhang, M. Chikaraishi, H. Seya, A. Fujiwara, "Effects of a GPS-enabled smart phone App with functions of driving safety diagnosis and warning information provision on overspeeding violation behaviour on expressways", Proceedings of the World Conference on Transport Research, Shanghai, China, 2016, pp. 1815-1823.

[14] J. Zhang, H. Lin, "Driving Behaviour Analysis and Traffic Improvement using Onboard Sensor Data and Geographic Information", Proceedings of the 7th International Conference on Vehicle Technology and Intelligent Transport Systems, 2021, pp. 284-291.

[15] H. Xiang, J. Zhu, G. Liang, Y. Shen, "Prediction of Dangerous Driving Behaviour Based on Vehicle Motion State and Passenger Feeling Using Cloud Model and Elman Neural Network", Frontiers in Neurorobotics, April 2021.

[16] S. Dixit, P. Thomas, C. Agarwal, "A Cloud-based Driver Monitoring for Inefficient Driving Behaviour using OBD2 Telematics", International Journal of Computer Applications. Vol. 153, No. 9, 2016.

[17] Y. Jeon, C. Kim, H. Lee, Y. Baek, "Real-time Aggressive Driving Detection System based on In-vehicle Information using LoRa Communication", Proceedings of the MATEC Web of Conferences 308, 2020.

[18] J. Meseguer, C. Calafate, J. Cano, DrivingStyles: "Assessing the Correlation of Driving Behaviour with Heart Rate Changes", Proceedings of the Smart Objects and Technologies for Social Good. Second International Conference, 2016.

[19] Y. He, S. Yang, X. Zhou, X.-Y. Lu, "An Individual Driving Behaviour Portrait Approach for Professional Driver of HDVs with Naturalistic Driving Data", Computational Intelligence and Neuroscience, Vol. 22, 2022.

[20] A. Agrawal, I. Gupta, "Rash Driving Detection Through Data Analytics", International Journal of Computer Sciences and Engineering, Vol. 7, No. 7, 2019.

[21] A. Sonawane, S. Dhabe, U. Nadgouda, V. Jadhav, G. V. Mane, "A Systematic Approach for Real-Time Clustering and Analysis of Drivers' Driving Behaviour using OBD-II", International Journal of Engineering Science and Computing, Vol. 6, No. 5, 2016, pp. 5047-5052.

[22] R. Massoud, F. Bellotti, R. Berta, A. De Gloria, S. Poslad, "Eco-driving Profiling and Behavioural Shifts Using IoT Vehicular Sensors Combined with Serious Games", Proceedings of the IEEE Conference on Games, 2019.

[23] N. S. Trindade, A. H. Kronbauer, H. G. Aragão, J. Campos, "Driver Rating: a mobile application to evaluate driver behaviour", South Florida Journal of Development, Vol. 2, No. 2, 2020, pp. 1147-1160.

[24] X. Fu, H. Meng, X. Wang, H. Yang, J. Wang, "A hybrid neural network for driving behaviour risk prediction based on distracted driving behaviour data", PLoS ONE, Vol. 17, No.1, 2022, p. e0263030.

[25] A. Mohammed, "Differences of Malaysian Bus Drivers Behaviours in Speeding, Acceleration and Deceleration Under Various Driving Conditions", International Journal of Mechanical Engineering, Vol. 7, No. 1, 2022.

[26] S. K. Alluhaibi, M. S. N. Al-Din, A. Moyaid, "Driver Behaviour Detection Techniques: A survey", International Journal of Applied Engineering Research, Vol. 13, No. 11, 2018, pp. 8856-8861.

[27] Z. Wang, F. Liu, X. Wang, Y. Du, "Driver Modeling Based on Vehicular Sensing Data", Proceedings of the International Conference on Advanced Control, Automation and Artificial Intelligence, Vol. 155, 2018, pp 137-141.

[28] E. Ostertagová, O. Ostertag, "Methodology and Application of One-way ANOVA", American Journal of Mechanical Engineering, Vol. 1, No. 7, 2013, pp. 256-261.

Human Face Recognition and Age Estimation with Machine Learning: A Critical Review and Future Perspective

Review Paper

Kavita
Department of Computer Science and Applications,
Maharshi Dayanand University, Rohtak, Haryana, India
kavitavce@gmail.com

Rajender Singh Chhillar
Department of Computer Science and Applications,
Maharshi Dayanand University, Rohtak, Haryana, India
chhillar02@gmail.com

Abstract *– Face Recognition (FR) applications are becoming more and more common these days. Face recognition, techniques, tools, and performance are all shown in this work, along with a literature review and gaps in many areas. Some of the most common uses of the FR include medical and government sectors as well as educational institutions. The FR technique can identify an appropriate individual through a camera. Online courses, online FDPs, and Webinars are becoming more interactive nowadays. Using Machine Learning, it is possible to quickly and securely determine a student's unique id to administer virtual online tests. The paper is an analysis of Machine learning and deep learning algorithms as well as tools such as Matlab and Python. The paper covers a survey of different aspects such as face detection, face recognition, face expressions, and age estimation. Hence, this is helpful for researchers to choose the right direction for their research. Future face recognition research is also considered in the paper which is now trending in face recognition systems. Data from recent years are used to evaluate the performance.*

Keywords*: Face Recognition, Face Expression, Age Estimation, Machine Learning, Python*

1. INTRODUCTION

Face recognition is a real-time application that has been considered a process of identifying individual faces by employing a distinct framework with different poses on datasets. [1] Nowadays, FR is an exciting field despite challenges also for researchers for decades. Facial recognition is the task of doing identification of a human face in a picture or video image for an existing database of faces. It starts with detecting different human faces from other images and after that, works on the identification of the detected image faces Face recognition is a tough task for identifying and verifying a person in a photograph; to detect a face from the number of faces with different expressions and emotions [2].

FR has a prominent role performed by humans in different circumstances such as photos captured under different effects of light and faces are changed with a change in obstructed or accessories like facial hair or change by age [3]. In FR, the age and gender of a human can also be predicted [4,5]. It remained challenging that faces a problem to predict older face images from a given children's face image, which might be very helpful to find the missing children after several years or presently he/she has in adult age. Commercial applications effectively operate faces that have the age of 18 yrs or greater than 18 [6]. Nowadays, iPhone considers the age to be 15 yrs. Therefore, age estimation also becomes challenging. Many researchers used various techniques to deal with children. For this purpose, a deep convolution network (CNN) was used to detect newborn babies' databases [7-10]. Face recognition has a wide range of applications, including access control, identity verification, educational institutions, businesses, banks, smartphones, security systems, surveillance systems, social media networks, and many more.

2. MACHINE LEARNING (ML)

ML has the subset of artificial intelligence (AI) where machines are trained by learning and improving without the interference of humans, adjust actions accordingly and themselves have the power of decision making. It endues the system the ability to work automatically and make decisions from experience except doing

any external programming [12]. It is the study of algorithms in computer science for accurate predictions and evaluation to behave intelligently. For this, Machines train to learn from prior knowledge existing in the knowledge base. Moreover; It is the building of programs that allows firstly to analyze data patterns and then classify patterns of images to find out behavior and make decisions for new input environments or uncertain situations in the real world. There is a various algorithm in Machine Learning that uses statistical techniques to predict patterns and then carry out actions on the patterns. The classification is processed based on past experiences and records. Three types of ML algorithms exist supervised, unsupervised, and reinforcement learning. For a new result, supervised ML algorithms are used for the labeled instances and known training datasets. For this, the first starts with the training dataset, and the learning algorithm produces an inferred function to estimate the predictions for the output data values. This system is learned with various new inputs and finds out the outcome and then the outcome is compared with the correct and desired outcome and errors are also find out to modify the system respectively so that the system becomes more powerful. Supervised learning is task-driven. Classification and Regression are solved with supervised learning and firstly; data is used for the training dataset and after that tested with a dataset that is similar or different to check the accuracy of the model. This algorithm is mainly used in Predictive Modeling. Popular Algorithms are Linear Regression, Logistic Regression, Naive Bays, Nearest Neighbor, DT, KNN, SVM, Convolution Neural Networks, and many more. Unsupervised Machine Learning algorithms apply to the unlabeled examples and information is not classified to train a model. In unsupervised learning, the algorithm produces an inferred function to represent a hidden structure. To draw a hidden structure; the system didn't evaluate the correct outcome; rather it can explore the data and can make interpretations with datasets from unlabeled data. This learning algorithm is mainly used in Descriptive Modeling. This learning is data-driven and used in Clustering, for Anomaly Detection. Popular algorithms are k-means and Association rules. Reinforcement Machine Learning algorithms are used to interact with the environment and search to find out the best possible path/outcome. Use the training dataset for the desired new result; learned from the experience and knowledge to produce correct decisions based on the received feedback. This is based on the hit & trial method and learns automatically with time variation. This algorithm is modeled as a Stock trading system, the Markov model and the most popular algorithms used in reinforcement learning are Q-Learning, Deep Networks, Computer games; Chess and GO, Self-driving cars, etc. [16-18].

3. FACE RECOGNITION

Face Recognition is a method of recognition that is capable to match the identity of an individual face from a given data set. There are so many methods of identification that may be more accurate, despite face recognition taking place an important role for researchers because of the non-interfering, powerful, and most natural way to detect a human face for confirmation of identification. Face recognition have prominent features in many areas such as identity verification in smart devices, Face Indexing, criminal investigations, diagnosing diseases, smart card applications, identifying people on a social platform, security camera, video surveillance, Face ID, Multimedia Environment, and many more. In the face recognition process, firstly input the image for preprocessing then recognition of the face is done by training the recognizer [19,21,22].

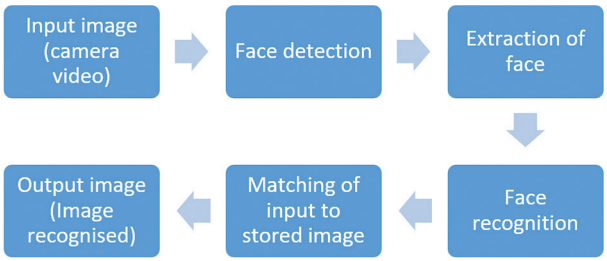


Fig. 1. Face Recognition Block Diagram

After teaching, testing, and the recognizer on the dataset(s) to achieve the results, there will be a positive outcome if the faces are correctly matched, else it will be negative. Retrain the face recognizer and test with the dataset again if a negative result is obtained. Face recognition may be done using a variety of algorithms. Any change in algorithm and dataset influences the accuracy of the face recognition system. The three steps of the Face Recognition process are:

Face Detection: faces are detected from the image or video. Apply preprocessing to know the exact location/coordinates of the face and extract that face for further processing then

Feature Extraction: by cropping the face from the image and extracting features from it. Here may use face embeddings to extract the features from the face.

Face Recognition: compare faces for embeddings of every face in the data saved in a file and classify whether the face of a person is matched or not.

In a neural network, a person's image begins as an input, and output in form of the vector is known as face embedding in the case of machine learning. To recognize a new image that is not in the dataset, the first step is to compute the face embedding for the image using the same network and then compare this embedding with the rest of the embeddings. To recognize the face; if the generated embedding is closer or similar to any other embedding were passed two images, as image1 did not save the embeddings but saved the embeddings of image2. Thus, when we compared the two new embeddings with the existing ones, the vector for image2 is closer to the other face embeddings of image2, whereas the face embeddings of image1 are not closer to any other embedding and thus the program cannot recognize the image1.

4. LITERATURE SURVEY

After the analysis of many papers, and research surveyed in the FR field, the face recognition surveys include face recognition, facial expressions, and estimation of age. This paper includes the latest papers. The study shows gaps for researchers. For the development of the face recognition technique; 23 datasets were used and have provided the essence of scholars' papers including methods, performance, and limitations. The performance might vary on different datasets [1]. The research focused on facial expressions and presented a systematic review. There was a study of techniques and algorithms of face detection that was proposed and abbreviated such as PCA, LBP, OF, Gabor filters, and many more for a better understanding. CK+ databases were used to detect difficulty with correctness during cognition of facial expressions with the environment. When poses were changed then system accuracy was suffered [2].

The research proposed a statistical model that detected measures of smiling faces. When a person smiles then the model measured different aspects of the face such as the mouth region, & geometric variations around facial parts, and also count active features during a smile. The proposed framework used 210 different parameters for smiling faces. For classification, SVM and the K-NN methods with CK+ and the MUG databases were used to classify gender. When measuring the smile of a person they assumed that; some discussions represent that females always have significant nature in expressing facial features than men and one of them is a smiling face. With the K-NN algorithm, the rate of classification was 86%, and to predict gender it was 85% achieved an overall rate of gender classification was 75% attained in this research [3]. The analysis of recent research in aging was also included. The estimate of age calculated by techniques such as MAE or CS and existing modeling techniques such as Linear Binary Pattern, Gabor Filter, Linear Discriminant Analysis, Principal Component Analysis, and Local Directional Pattern has been used. Then the performance of various systems was compared and evaluated to estimate the age of people who were challenged with the number of datasets such as FGNet, Morph, and Gallagher Collection Person datasets that are public databases. Morph, NIS, YGA, LHI image database, and Gallagher Collection Person databases are large databases and more compatible to evaluate the gender and age of a person with the regression technique [4].

After that; a comparative study of OpenCV and "dlib" were performed and further observed time complexity. Therefore; combine phases to make the system better performance rather than a single one because OpenCV was more effective and performed better results than the "dlib" library in face identification and recognition applications for the IoT platform as well. When combined LBPH with HOG then the results of facial recognition are better than the above algorithms such as Eigenface, and Fisher Face [5-7]. Unfiltered photos of faces, as well as age and gender groups, may all be predicted with reasonable accuracy using a convolution neural network. Two layers are employed to determine an individual's age or gender, a feature extraction layer, and a classification layer. The convolution network was already trained with the IMDB-wiki dataset and was on Morph_II images and at last with the OUI-adience dataset. The model results better with 16.6 classifications of age and achieves 3.0% more accuracy for the classification of gender [8].

Other techniques include such as Eigenfaces, SVM, and FaceNet neural networks on Facial Embedding. The facial embeddings have been set up by passing through a Pre-trained Network. Viola-Jones algorithm is used for face detection. The maximum accuracy obtained to classify gender is 97% and KNN has a better result than Logistic Regression, SVM, Naïve-Bayes, and Decision Trees [9]. With the addition of two normalization processes on two levels, CNN's modified architecture was further developed. The proposed algorithm applied batch normalization for the first as well as final convolution layers and achieved higher accuracy rates. Distinct facial traits were extracted & the Softmax classifier was employed in CNN's fully connected layer to categorize faces. It observed that the proposed approach on the Face dataset gave better results which improved the performance of face recognition [10]. A CNN system for recognition & identification of expression of a student's face was presented. The network has 4 levels in which 2 layers were fully connected and 2 were max-pooling layers finally created a model of facial emotion recognition that identifies the emotion of students from faces. The model has three methods: Haar Cascade detector to detect faces, normalization, and emotion recognition in Convolution network with FER dataset. The proposed system achieved a 70% rate of accuracy at 106 epochs. is helpful in education that helps teachers to recognize the interest of students by their expressions of faces during their studies. Used ResNet for facial emotion recognition, solve many problems like gradient points, and gave visual perceived results and better performance with the network system. The system obtained high accuracy i.e., 85.76 and 64.40 on training and testing datasets during convolution neural network and also do better with the FER database [11-14].

Face detection, recognition, and emotion classification of the face were classified as three phases: face detection, recognition, and emotion classification of the face. OpenCV and python were suggested for the CV technique. The expressions are categorized into seven states, each of which expresses a particular facial scenario. The Viola-Jones method was utilized for emotion recognition on the FER2013 dataset and an accuracy of 88% was achieved with happiness in images and the lowest accuracy with sadness was 57% detected. Hence the designed system performs better than existing ones and is helpful to analyze the emotions of

students' E-learning techniques [15]. The optimization was performed for seven expressions on the faces of persons. To identify the emotions convolution neural network, an approach with Keras and theano libraries was implemented for facial expressions. The Viola-Jones method was utilized for emotion recognition on the FER 2013 dataset [16].

A state-of-art model designed using python and supervised learning that might be helpful to detect emotions of the face that was captured with a web camera with 96% prominent accuracy. In addition, end-to-end CNN architecture was introduced. CNN model was trained for supervised face bounding boxes and personal identities using Wider_FACE and CasiaWeb FACE databases and tested with face detector and "fddb" and "lfw" data sets. The feature map was created using a spatial transformer network (STN) without the need for face alignment. The results come with 89% accuracy in the detection phase and 98% in the recognition phase. Therefore, it is better to conclude that rather than two models, combining them as a single one has better performance. After that, a system was used that might detect a person directly from a group of persons. For this, they used a deep convolution network that trained the model and gave good results by applying a filter to detect the face of a person from multiple gestures [17-18]. Stochastic gradient descent (SGD) was used to train the network using the Celeb Faces attribute dataset (Celeb A) and achieved 99.7% accuracy on the Labeled wild_ Faces and 94% on YTF databases [19-21].

4.1 RESEARCH GAPS, FUTURE SCOPES

Research always explores a path for further research. Therefore firstly, consider a deep study of these papers and then find out the limitations/gaps in the papers and also find out the future scope for further research. Research gaps and future scopes along with techniques and limitations are discussed in table 1.

Table 1. Research Gaps and Future Scopes

Citation	Publication, year	Title of Paper	Techniques used	Research Gaps/ Limitations	Future Scope
[1]	IEEE, 2018	Changing Facial Features in Face Recognition and Age Estimation: A Critical Review Study	SVM, LBP, GAP	The system face problem with the evaluation of age effects.	The system can be implemented with newborn child's faces.
[22]	ACM, 2018	Face recognition: Sparse Representation vs. DeepLearning	SRC, CNN, Multi-PIE and YTC databases	Noted that SRC has small databases. Errors occurred during recognition of gestures &, expression of the face was challenged.	Face recognition might be further rectified with the merging of CNN locality & SRC linearity to enhanced variants.
[15]	IEEE, 2019	An investigation of the effectiveness of facial recognition systems on humans	Eigenfaces, Fisherfaces, and LBPH algorithm	Many faces were identified with a single data when using the LHBP database.	In the future, can elaborate with ambient analysis & implementation for the recognition of faces with distinct angles & poses.
[8]	Elsevier, 2019	Facial Embeddings for Gender Classification: A New Approach	KNN, Facenet NetworkViola-Jones algorithm, python	In the UTK Face dataset, only 203 images for training and testing; work with a small database.	This work can be explored to obtain better results with the estimation of age by including race value, and place of the center of a person.
[18]	IEEE, 2019	Student Emotion Detection with CNN's Facial Emotion Recognition	Haar Cascades, CNN on FER 2013 dataset and OpenCV library.	The system got confused in fear & sad faces with feared expression results were comparatively poor.	In the future, can try to apply the CNN framework with 3D face images of students with facial expressions.
[12]	Science Direct, 2019	Imperfect facial data can be used to train deep face recognition systems.	CNN, VGGF, and SVMs as well as CS models for facial feature extraction.	the dataset utilized was far excluded from the practical session and work was on data which incorrect in the past.	Extend this work with Cctvcameraimages for face detection.
[6]	Hindawi, 2020	For Age and Gender Predictions of Unfiltered Faces, Deeply Applied Classifiers	two-level CNN Model, robust image preprocessing algorithm	Challenging conditions with resolution, lightning effects, and deep makeup on OUI_ imagesdataset.	A deep CNN model with a lasting algorithm will be considered. Research can investigate the apparent age estimation approach.
[4]	Hindawi, 2020	Improved Mask R-CNN-Based Face Detection and Segmentation	FCN, faster R-CNN, Mask R-CNN, G-Mask model add Segmentation branch	Due to G- mask segmentation, complexity increased.	Try to improve the speed of the G-Mask method.
[5]	Elsevier, 2020	Facial emotion identification using a deep self-attention network	Deep learning framework with CNN, ResNet. FER dataset.	Work was restricted to only Y-channel. X-channel not covered inYcbcr for resized images	The system can be enhanced with other techniques of computer_vision using deep learning.

Citation	Publication, year	Title of Paper	Techniques used	Research Gaps/ Limitations	Future Scope
[32]	IEEE, 2020	Prediction of Emotions Using Facial Expression Recognition	CNN classifier and Csv image format were used.	The model suffered from fitting when emotions were detected in between 70 and 80 passes and then the model seems to overtrain.	This model extended to identify mood swings of the person due to situations that come in the environment which was a reason for varying behavior.
[38]	Springer, 2020	Face-recognition technology and gender bias in criminality detection	Deep network with SNN & CNN designed.	The only small size of dataset used, a problem arises when enlarging the dataset, occurred difficulty in measuring the impact of racial bias and subjectivity of identifying race	CNN may use to train and test enlarging datasets and also work in measuring the effect of race and gender bias to detect and express individuality signs.
[11]	Hindawi, 2020	FaceFilter: DL and Filter Algorithms for Face Recognition	Deep convolutional network. FaceNet, SGD algorithm.	The problem comes with many faces while the model detected faces correctly. This restriction shows poor results.	This work tries to make an adequate system that overcomes the limit of detection of faces.
[25]	IJERT, 2021	Real-Time Face Mask Detection & Recognition using Python	Face Recognition System, Python	There is a need to improve the accuracy	Research would be useful to find the face mask while entering a public location
[33]	Hindawi, 2021	Analysis & Implementation of Optimization Techniques for Facial Recognition	Face Recognition System	Need to do more work on performance enhancement	Research might be extended to improve scalability
[35]	ASTROS, 2021	Performance Evaluation of CNN And VGG on Real-Time Face Recognition	CNN, VGG, Face Recognition System	There is a need to integrate a compression mechanism to improve performance	Such research would play a significant role in image processing
[41]	Hindawi, 2021	CVS-based algorithm for digital image face detection and recognition	Face Detection	Research has limited scope and flexibility	Face detection and recognition enhancement could play a significant role in security systems.
[42]	Springer, 2022	Novel Face Detection Algorithm Based on Fuzzy Distance-Based Minimum Spanning Tree Clustering	Face Detection	Fuzzy logic and clustering mechanism implementation are found complicated	Research would lay a strong foundation of decision making in future research works
[43]	Springer, 2022	ML techniques for facial detection and recognition: a survey	Face Detection, Machine Learning	There is a lack of technical work	Such research would play a significant role in decision-making using a machine learning technique

5. PROBLEM STATEMENT

There have been several types of research in the area of face recognition. These techniques are considered image processing and machine learning approach to perform face recognition operations. But the issue with previous research is performance and accuracy. There is a need to do more work in the area of the face recognition system.

6. PROPOSED RESEARCH METHODOLOGY

In the proposed work research related to face recognition has been considered and the methodologies used in those researches are considered. The face recognition surveys also include facial expressions and estimation of age. This paper includes the latest papers for the year 2018-2022. The research layout certain difficulties faced in the face recognition area with facial features and prediction of age. Hence, the paper enlightens them with future scopes.

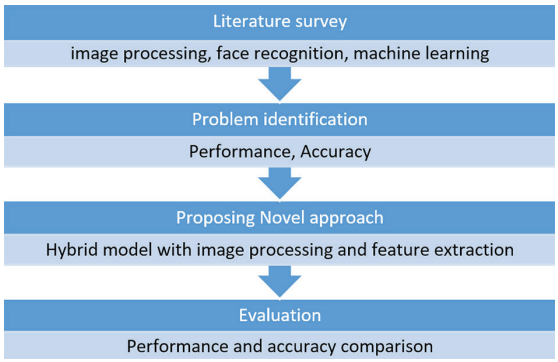


Fig. 2. Process flow of work

7. RESULT AND DISCUSSION

The performance of face recognition shown in table 2 is in terms of accuracy and demonstrates with a comparative study of algorithms for face detection. The model designed for face detection as well as recognition is summarized with different datasets.

7.1 ACCURACY OF FACE RECOGNITION SYSTEM

The results of the study show that higher the accuracy rate for KNN has 97% on a self-small dataset, for SVM has 95%, and LBPH at 80% on a self dataset. The CNN model has an accuracy rate was detected 83%, and with deep CNN of 99.7% on the LFW dataset.

Table 2. Performance of ML models with datasets

Citation	Author	Year	Algorithm	Machine Learning Model(s)	Dataset(s)	Accuracy(%)
[16]	Hassan et al.	2018	-	KNN, SVM	CK+, MUG	86
[17]	Lu et al.	2018	-	Deep coupled ResNet	LFW, face	99
[29]	G. P. Nam et al.	2018	-	PSI-CNN	LFW, CCTV	88.7
[8]	Avinash et al.	2019	Viola John	KNN	SELF	97
[12]	Ali et al.	2019	-	VGG (face)	FEI	70-90
[13]	Khan et al.	2019	-	Deep CNN	-	98.5
[15]	Sharmila et al.	2019	Haar Cascade	LBPH	SELF	80
[18]	Imane et al.	2019	Haar Cascade	CNN	FER 2013	70
[31]	Chen Qin et al.	2019	-	Deep CNN	-	94.6
[40]	Ishan et al.	2019	-	MatConv Net	SELF	94.8
[5]	Arpita et al.	2020	-	YCbCr	FER	64
[6]	Olatunbosun et al.	2020	-	CNN	IMDb-WIKI, MORPH-II	83
[7]	Shaik et al.	2020	Viola John	CNN	FER 2013	70
[9]	Hongxin et al.	2020	-	CNN + STN	FDDB, LFW	86.3
[10]	Ekaterina et al.	2020	Viola John	CNN	FER 2013	69
[11]	Mohammed et al.	2020	-	deep CNN	LFW	99.7
[32]	Kaustav et al.	2020		CNN	CSV	69
[33]	Justice et al.	2021	YouTu, PSO	SVM	AT&T	82.05
[41]	Di Lu et al.	2021	Seetaface	SVM	Yale-B	66
			BscGc	SVM, OpenCV	ORL	95

7.2 PERFORMANCE GRAPH OF DATASETS

The below graph provides information on recognition rate with measures of accuracy in the percentage of recent years with different datasets. The accuracy of a model varies with datasets, as shown in Fig. 3. The performance graph is taken in between various datasets used in the number of years.

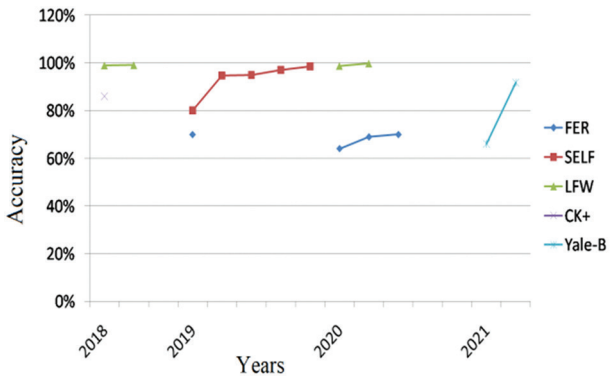


Fig. 3. Performance Graph of Datasets

8. CONCLUSION

Face recognition is a very promising field in today's era. The results of the study show that the deep CNN model has high accuracy with the LFW dataset compared to other datasets for FR. The accuracy rate for age estimation is 92% in the deep CNN model. This paper finds that the facial technique's performance was distinct for different data sets i.e. varied from one data set to another. The paper survey concludes that sometimes the system confuses with angry/sad and fear/surprise facial expressions, and challenges have come in age estimation also. Despite this, the paper focuses on future perspectives. Hence, a machine learning model can be designed to identify a person and predict the emotions of a face by using supervised learning more effectively to solve the real scenario challenges to enhance the performance of the system. After that work can be enhanced for age estimation as well.

9. REFERENCES

[1] R. R. Atallah, A. Kamsin, M. A. Ismail, S. A. Abdelrahman, S. Zerdoumi, "Face Recognition and Age Estimation Implications of Changes in Facial Features: A Critical Review Study", IEEE Access, Vol. 6, No. C, 2018, pp. 28290-28304.

[2] D. Canedo and A. J. R. Neves, "Facial expression recognition using computer vision: A systematic review", Applied Sciences, Vol. 9, No. 21, 2019, pp. 1-31.

[3] S. F. Kak, F. Mahmood, M. Alfiky, "A Review of Person Recognition Based on Face Model", Eurasian Journal of Science Engineering and Technology, Vol. 4, No. 1, 2018.

[4] R. Angulo, J. R. Tapamo, O. A. Aderemi, "Age estimation via face images: a survey", Journal on Image and Video Processing, Vol. 42, 2018, p. 42.

[5] K. Lin, H. Zhao, J. Lv, C. Li, X. Liu, R. Chen, R. Zhao, "Face Detection and Segmentation Based on Improved Mask R-CNN", Discrete Dynamics in Nature and Society, Vol. 2020, 2020, pp. 1-11.

[6] N. Boyko, O. Basystiuk, N. Shakhovska, "Performance Evaluation and Comparison of Software for Face Recognition, Based on Dlib and Opencv Library", Proceedings of the IEEE Second International Conference on Data Stream Mining & Processing, Lviv, Ukraine, 21-25 August 2018, pp. 478-482.

[7] S. Khan, M. H. Javed, E. Ahmed, S. A. A. Shah and S. U. Ali, "Facial Recognition using Convolutional Neural Networks and Implementation on Smart Glasses", Proceedings of the International Conference on Information Science and Communication Technology, Karachi, Pakistan, 9-10 March 2019, pp. 1-6.

[8] A. Olatunbosun, V. Serestina, "Deeply Learned Classifiers for Age and Gender Predictions of Unfiltered Faces", The Scientific World Journal, 2020, p. 12.

[9] A. Swaminathan, M. Chaba, D. K. Sharma, Y. Chaba, "Gender Classification using Facial Embeddings: A Novel Approach", Procedia Computer Science, Vol. 167, 2020, pp. 2634-2642.

[10] I. R. Pandey, M. Raj, K. K. Sah, T. Mathew, M. S. Padmini, "Face Recognition Using Machine Learning", International Journal of Science &Technoogy, Vol. 6, No. 4, 2019.

[11] A. Gupta, S. Arunachalam, R. Balakrishnan, "Deep self-attention network for facial emotion recognition", Procedia Computer Science, Vol. 171, 2020, pp. 1527-1534.

[12] I. Lasri, A. R. Solh, M. E. Belkacemi, "Facial Emotion Recognition of Students using Convolutional Neural Network", Proceedings of the Third International Conference on Intelligent Computing in Data Sciences, Marrakech, Morocco, 28-30 October 2019, pp. 1-6.

[13] A. Elmahmudi, H. Ugail, "Deep face recognition using imperfect facial data", Future Generation Computer Systems, Vol. 99, 2019, pp. 213-225.

[14] Z. Lu, X. Jiang, A. Kot, "Deep Coupled ResNet for Low-Resolution Face Recognition", IEEE Signal Processing Letters, Vol. 25, No. 4, 2018, pp. 526-530.

[15] S. A. Hussain, A. Salim Abdallah Al Balushi, "A real time face emotion classification and recognition using deep learning model", Journal of Physics: Conference Series, Vol. 1432, No. 1, 2020.

[16] E. Ivanova, G. Borzunov, "Optimization of machine learning algorithm of emotion recognition in terms of human facial expressions", Procedia Computer Science, Vol. 169, 2020, pp. 244-248.

[17] K. D. Aljafaar, "Face recognition using Viola-Jones depending on Python", Indonesian Journal of Electrical Engineering and Computer Science, 2020, pp. 1513-1521.

[18] S. M. Bah, F. Ming, "An improved face recognition algorithm and its application in attendance management system", Array, Vol. 5, 2020, p. 100014.

[19] H. Zhang, L. Chi, "End-to-End Spatial Transform Face Detection and Recognition", Virtual Reality & Intelligent Hardware, Vol. 2, No. 2, 2020, pp. 119-131.

[20] K. M. Pathak, S. Yadav, P. Jain, P. Tanwar, B. Kumar, "A Facial Expression Recognition System To Predict Emotions", Proceedings of the International Conference on Intelligent Engineering and Management, London, UK, 17-19 June 2020, pp. 414-419.

[21] M. Alghaili, Z. Li, H. A. R. Ali, "FaceFilter: Face Identification with Deep Learning and Filter Algorithm", Scientific Programming, Vol. 2020, 2020.

[22] N. H. Alskeini, K. N. Thanh, V. Chandran, W. Boles, "Face recognition: Sparse representation vs. Deep learning", The ACM International Conference Proceeding Series, October 2018, pp. 31-37.

[23] Sharmila, R. Sharma, D. Kumar, V. Puranik, K. Gautham, "Performance Analysis of Human Face Recognition Techniques", Proceedings of the 4th International Conference on Internet of Things: Smart Innovation and Usages, Ghaziabad, India, 18-19 April 2019, pp. 1-4.

[24] M. Khan, S. Chakraborty, R. Astya, S. Khepra, "Face Detection and Recognition Using OpenCV", Proceedings of the International Conference on Computing, Communication, and Intelligent Systems, 2019, pp. 116-119.

[25] R. M. Thomas, M. Sabu, T. Samson, S. Mol, T. Thomas, "Real Time Face Mask Detection and Recogni-

tion using Python", International Journal of Engineering Research & Technology, Vol. 9, No. 7, 2021, pp. 57-62.

[26] L. Li, X. Mu, S. Li, H. Peng, "A Review of Face Recognition Technology", IEEE Access, Vol. 8, 2020, pp. 139110-139120.

[27] S. Almabdy, L. Elrefaei, "Deep Convolutional Neural Network-Based Approaches for Face Recognition", Applied Sciences, Vol. 9, No. 20, 2019, p. 4397.

[28] P. S. Prasad, R. Pathak, V. K. Gunjan, H. V. Ramana Rao, "Deep Learning Based Representation for Face Recognition", Lecture Notes in Electrical Engineering, Vol. 570, 2020, pp. 419-424.

[29] G. P. Nam, H. Choi, J. Cho, I. J. Kim, "PSI-CNN: A Pyramid-based scale-invariant cnn architecture for face recognition robust to various image resolutions", Applied Sciences, Vol. 8, No. 9, 2018.

[30] H. Ugail, A. Al-dahoud, "Is gender encoded in the smile? A computational framework for the analysis of the smile driven dynamic face for gender recognition", Visual Computer, Vol. 34, 2018, pp. 1243-1254.

[31] C. Qin, X. Lu, P. Zhang, H. Xie, W. Zeng, "Identity Recognition Based on Face Image", Journal of Physics: Conference Series, Vol. 1302, No. 3, 2019.

[32] M. Hashemi, M. Hall, "Criminal tendency detection from facial images and the gender bias effect", Journal of Big Data, Vol. 7, No. 1, 2020.

[33] J. K. Appati, H. Abu, E. Owusu, K. Darkwah, "Analysis and Implementation of Optimization Techniques for Facial Recognition", Applied Computational Intelligence and Soft Computing, Vol. 2021, 2021, 13 pages.

[34] U. Chavan, D. Kulkarni, "Optimizing deep convolutional neural network for facial expression recognitions", Advances in Intelligent Systems and Computing, Vol. 808, No. 2, 2019, pp. 185-196.

[35] S.A. Dar, S. Palanivel "Performance Evaluation of Convolutional Neural Networks (CNNs) And VGG on Real Time Face Recognition System", Advances in Science, Technology and Engineering Systems Journal, Vol. 6, No. 2, 2021, pp. 956-964.

[36] M. K. Hasan, M. S. Ahsan, Abdullah-Al-Mamun, S. H. S. Newaz, G. M. Lee, "Human face detection techniques: A comprehensive review and future research directions", Electronics, Vol. 10, No. 19, 2021.

[37] N. Gupta, P. Sharma, V. Deep, V. K. Shukla, "Automated Attendance System Using OpenCV", Proceedings of the 8th International Conference on Reliability, Infocom Technologies and Optimization, Noida, India, 4-5 June 2020, pp. 1226-1230.

[38] Y. Shi, A. Jain, "Probabilistic face embeddings", Proceedings of the IEEE/CVF International Conference on Computer Vision, Seoul, Korea, 2019, pp. 6901-6910.

[39] Z. Yuan, "Face Detection and Recognition Based on Visual Attention Mechanism Guidance Model in Unrestricted Posture", Scientific Programming, Vol. 2020, 2020.

[40] K. M. Rajesh, M. Naveenkumar, "A robust method for face recognition and face emotion detection system using support vector machines", Proceedings of the International Conference on Electrical, Electronics, Communication, Computer and Optimization Techniques, Mysuru, India, 9-10 December 2016, pp. 1-5.

[41] D. Lu, L. Yan, "Face Detection and Recognition Algorithm in Digital Image Based on Computer Vision Sensor", Journal of Sensors, Vol. 2021, 2021, 16 pages.

[42] Y. Li, W. Zhou, "A Novel Fuzzy Distance-Based Minimum Spanning Tree Clustering Algorithm for Face Detection", Cognitive Computation, 2022, pp. 1350-1361.

[43] D. Gupta, A. Khanna, V. Kansal, G. Fortino, A. E. Hassanien, "Advances in Intelligent Systems and Computing", Proceedings of Second Doctoral Symposium on Computational Intelligence, Vol 1374, Singapore, 2022.

Software Defect Prediction using Deep Learning by Correlation Clustering of Testing Metrics

Original Scientific Paper

Kamal Kant Sharma
Department of Information Technology, KIET Group of Institutions, Delhi-NCR, Ghaziabad
Dr. A.P.J. Abdul Kalam Technical University, Lucknow, India
knmietkamal@gmail.com

Amit Sinha
Department of Information Technology, ABES Engineering College, Ghaziabad
Dr. A.P.J. Abdul Kalam Technical University, Lucknow, India
amit.sinha@abes.ac.in

Arun Sharma
Department of AI and Data Sciences
Indira Gandhi Delhi Technical University for Women, Delhi, India
arunsharma2303@gmail.com

Abstract *– The software industry has made significant efforts in recent years to enhance software quality in businesses. The use of proactively defect prediction in the software will assist programmers and white box testing in detecting issues early, saving time and money. Conventional software defect prediction methods focus on traditional source code metrics such as code complexities, lines of code, and so on. These capabilities, unfortunately, are unable to retrieve the semantics of source code. In this paper, we have presented a novel Correlation Clustering fine-tuned CNN (CCFT-CNN) model based on testing Metrics. CCFT-CNN can predict the regions of source code that contain faults, errors, and bugs. Abstract Syntax Tree (AST) tokens are extracted as testing Metrics vectors from the source code. The correlation among AST testing Metrics is performed and clustered as a more relevant feature vector and fed into Convolutional Neural Network (CNN). Then, to enhance the accuracy of defect prediction, fine-tuning of the CNN model is performed by applying hyperparameters. The result analysis is performed on the PROMISE dataset that contains samples of open-source Java applications such as Camel Dataset, Jedit dataset, Poi dataset, Synapse dataset, Xerces dataset, and Xalan dataset. The result findings show that the CCFT-CNN model increases the average F-measure by 2% when compared to the baseline model.*

Keywords*: Software Engineering, Software Testing, Abstract Syntax Tree, Machine Learning, Convolution Neural Network*

1. INTRODUCTION

Software engineering is a vast field that is extensively used in software-based industrial applications. Humans have become increasingly reliant on software-based applications in subsequent years, with software functionality being regarded as the most important aspect of customer experience [1]. Numerous consequences, like a breakdown of software functioning and incorrect output, could arise during real-time software-based applications, resulting in consumer dissatisfaction. Nonetheless, when a large number of activities are processed and performed by software, the software's performance may be reduced. Real-time applications frequently experience software failures and flaws. Program error appears when performance differs from predicted results, causing disagreement between real and necessary software functioning, often known as a software flaw or bug [2]. As the need for user apps grows, so do the complexities of software applications, which has an impact on the operations of numerous software corporations and leads to the production of inadequate software uses. Due to a large number of software applications, experts find it difficult to identify, locate, and discover software defects. Moreover, in the field of software fault prediction and diagnosis, defect density is a challenging problem. Previous research and expertise in this field can help forecast software product faults. Initially, manual testing was used to uncover software defects in the software industry. In the overall development of a program, manual software testing demands 27% human labor [3]. Manual testing takes more time and labor, and it can't catch all of the faults in the program. Defect forecasting models are commonly employed by companies to solve this problem. During the development process, these models allow firms in fault forecasting, effort estimate, software reliability testing, risk assessment, and other functions [4]. This can also contribute to minimizing risk by utilizing software quality forecasting models early in

the software development lifecycle (SDLC), resulting in increased customer contentment and cost reductions, as well as reduced human work. To solve the problem of software fault forecasting, many strategies have been developed. In this context, software metrics are useful for forecasting software product problems by examining the link between software metrics and output product reliability. Process metrics, project metrics, and product metrics are the three types of software metrics. Software development and maintenance utilize process metrics to ensure the software's operation and longevity. Product metrics explain different attributes like layout features, quality level necessity, and performance level need for any specific software program, whereas project metrics offer various parameters including the total number of developers, developer expertise, work scheduling, organization, and software size [5].

Machine learning and data mining approaches are regarded to be the most effective in the field of software engineering and fault predictions. Machine learning is highly utilized in a variety of implementations due to its potential to meet or surpass human-level functionality. The accessibility of huge data samples, high-speed processing techniques, and their capabilities to accomplish excellent functionality has developed the way for DNNs to enter safety-critical applications which include automated vehicle driving, healthcare diagnostic test, safety, and so on. Because such implementations are safety-critical, sufficient testing of these machine learning is required before implementation. Nevertheless, unlike conventional software, machine learning lacks a clear control-flow structure. They learn their judgment call strategy by instructing on a big dataset and progressively modifying variables utilizing various means to accomplish the required precision. As a result, conventional software checking techniques such as operational coverage, branch coverage, and so forth cannot be implemented in machine learning, posing a challenge to their usage in safety-critical implementations [1].

Conventional software screening techniques perform badly when implemented in machine learning since machine learning code contains no details about a machine learning internal decision-making rationality. Machine learning comprehends its regulations from instructing information and does not have the control-flow structure found in conventional software [3]. As a result, conventional coverage criteria such as code coverage, branch coverage, functional coverage, and so on are inapplicable to machine learning. Fig 1 depicts a high-level depiction of the majority of established Whitebox evaluation techniques for machine learning. The machine learning, evaluation inputs, and coverage metric are used as inputs to the evaluation procedure to guarantee that all sections of the program logic have been examined. An oracle determines whether the machine learning action is accurate for the examined inputs. A coordinated test input generation technique can also be utilized to produce test input information with wider reach and expose more corner case behavioral patterns. Established analysis techniques typically produce either an estimate of system appropriateness or an adversarial ratio [3].

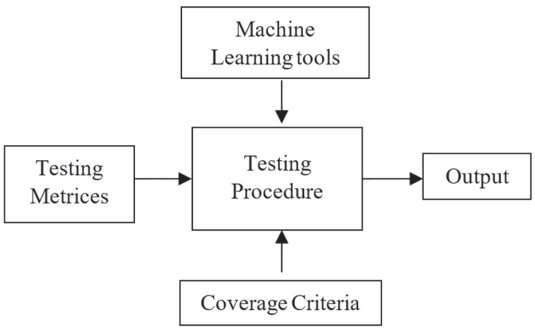


Fig 1. Representation of Testing using Machine Learning

Machine learning (ML) systems are highly being used in safety-critical areas as a result of subsequent advancements in the domains [4]-[10]. As a result, several software analysis techniques are necessary to guarantee the dependability and security of ML systems. Because ML framework regulations are interpreted from training information, it is challenging to understand the execution regulations for every ML system behavior. At the same time, Random Testing (RT) is a prominent testing technique, and it does not require any understanding of software execution. As a result, RT is an excellent choice for evaluating ML systems. Furthermore, the established methods for evaluating ML frameworks rely highly on RT by labeled testing datasets. Random Testing (RT), being one of the most basic and widely used software screening techniques, could be used to evaluate ML systems. The advantage of RT is that it is simple in theory and convenient to apply [5]. Most notably, if researchers have no understanding of software execution, it could be the only practicable method [6]. Undoubtedly, due to the change in advancement framework caused by ML, this feature is also an essential advantage for evaluating ML systems [7]. Conventional software is built axiomatically by attempting to write down the regulations as program codes, and the system's behavior is controlled by these known regulations. The execution guidelines for ML systems, on the other hand, are derived inductively from retraining information and are not precisely known to us. Because of the paradigm shift, several of the key characteristics of ML systems have no source code. As a result, RT is an excellent evaluation technique for ML systems. In reality, the conventional way for evaluating ML systems is to collect and physically categorize as much test information as possible, then utilize this data to evaluate ML systems using the RT technique [8].

2. RELATED WORK

Gerasimou et al. [8] focused on altering typical software engineering assessment requirements to increase certainty in their right conduct. Nevertheless, they fall

short of representing the essential features that these systems display. Wang et al [9] presented Robustness-Oriented Assessment, a unique assessment methodology (RobOT). A major component of RobOT is a quantitative assessment of 1) the utility of every testing ground in enhancing model durability (typically via additional training) and 2) the model resilience improvement's converging reliability. Guo et al [10] presented Audee, a unique method for evaluating DL systems and locating flaws. Audee uses a search-based method and three distinct mutation algorithms to produce a variety of testing cases by combining model architectures, variables, evaluations, and inputs. Audee can identify three different kinds of problems: logical flaws, failures, and Not-a-Number (NaN) issues. Ma et al [11] presented a mutation assessment method for DL systems to evaluate the integrity of testing information. To achieve so, they provided a collection of source-level mutation operators to introduce flaws into the sources of DL, in a similar manner as mutation assessment in conventional software (i.e., training data and training programs). Ma et al [12] presented DeepGauge, a collection of multi-granularity evaluating requirements for DL systems that tries to generate a multi-faceted depiction of the testing ground, in this study. On two well-known databases, five DL systems, and four state-of-the-art confrontational threat approaches against DL, the in-depth examination of their recommended assessment requirements is proven. DeepGauge's potential utility offers insight into the development of more comprehensive and powerful DL systems. Du et al [13] took the initial move toward evaluating RNN-based stateful DL systems in this research. They characterized RNN as an abstract state transition system, from which they derived a series of stateful DL-specific testing coverage requirements. Wang et al [14] presented a new method for detecting challenging specimens in real-time. When they applied mutations to the DNN, they found that adversarial specimens are substantially more sensitive than regular specimens. They initially constructed a vulnerability metric. CBIL is the name of the hybrid model that was proposed by Farid et al. [15]. CBIL can identify problematic parts of source code. It does it by pulling vector representations of Abstract Syntax Tree (AST) tokens out of the source code. The transformation of integer vectors into dense vectors can be accomplished through mapping and word embedding. The AST tokens' semantics are then extracted by a CNN, which stands for the convolutional neural network. After then, the Bidirectional Long Short-Term Memory (Bi-LSTM) system will retain the most important features while ignoring the less important features to improve the accuracy of software defect prediction. Fan et al. [16] developed a hierarchical method to automatically evaluate the degree of similarity between mapped statements and tokens using a variety of distinct algorithms. The author was able to identify whether or not each of the compared algorithms provides erroneous mappings for a sentence or the tokens that make up the claim by doing the comparison. The study of the results reveals that the precision is between 0.98 and 1.00, and the recall is between 0.65 and 0.75.

3. METHODOLOGY

Human error is the primary cause of software bugs. There will be a defect (fault, bug) in the code because of these mistakes. There is no bug-free software. The majority of applications are classed as Critical, High, Medium, or Low because of the number of flaws they include. Problems of all sizes can be caused by the Metrics used.

Defect prediction is used to identify the sections of software that are most likely to cause problems in future versions. Choosing the appropriate testing Metrics is the first process. Then, if there are any errors in the data, it is marked as defective. Otherwise, it is marked as clean. The key features or testing Metrics of each software code file are extracted and collected. This step is critical. Finally, the model is constructed and tested using the labeled data and collected features. Finally, machine learning is used to determine whether or not the new instance is faulty or otherwise clean.

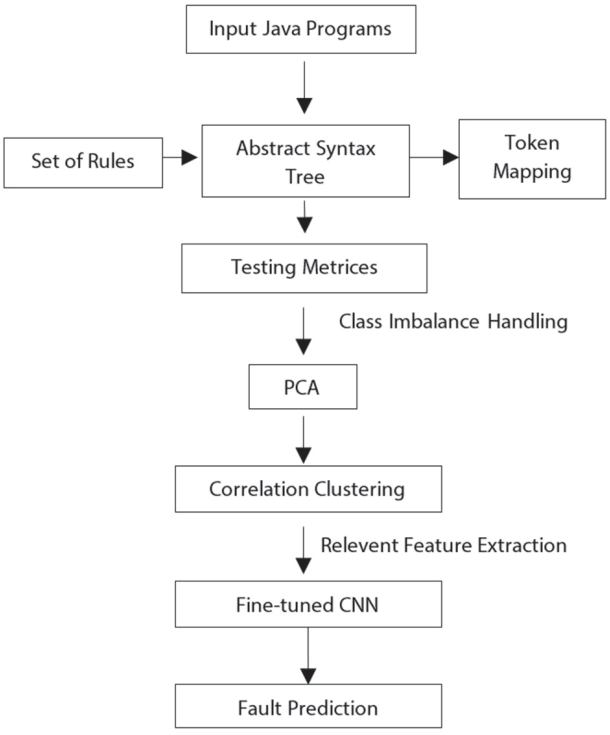


Fig 2. Flowchart of Methodology

In the model, there are two datasets: the training set is used to create and train the model, and the test set is used to evaluate and evaluate the trained model. Software defect prediction models can be divided into two categories. Within-Project Defect Prediction (WPDP) and Cross-Project Defect Prediction (CPDP) are two types of defect prediction. As a result of this, WPDP uses data from previous editions of the project. The same project is used for both the training and testing sets.

Two separate projects are used to train and evaluate the application in the CPDP, with the training set coming from one project and the test set coming from a separate platform. In this paper, we have adopted the CPDP model for performance evaluation.

In this paper, cascaded correlation clustering fine-tuned CNN model is proposed and termed (CCFTCNN). Fig. 2 depicts the complete framework for the proposed software testing module using machine learning. In this paper, testing Metrics are extracted using Abstract Syntax Tree (AST) for all java sources in the learning and testing groups. The vector of textual characters is then transformed into a numeric vector by creating mapping dictionaries between symbols and numbers. Then, text encoding is used to convert every numeric vector into a dense vector. Finally, the resulting dense vector is utilized as information for developing and training the proposed framework i.e., CCFTCNN. Then its functionality is examined using the PROMISE database.

3.1 DATA PREPROCESSING

In this stage, data is pre-processed before learning to increase its correctness and address the class balancing problem, which might have an adverse effect on the outcomes of the framework.

3.2 ABSTRACT SYNTAX TREE (AST)

Initially, an Abstract Syntax Tree (AST) is developed for the chosen Java items in the PROMISE database. Furthermore, the types of AST nodes that will be created as symbols are as follows: (1). Control flows elements, like if/while/do words, which are documented as node kinds. (2). nodes of class example constructions and technique incantation that are reported without parenthetical as the titles of types or techniques (3). proclamation elements, including technique/categories/ declarations, whose titles are marked. Other AST nodes are deleted because they can have an impact on the significance of the AST nodes that have been chosen. Figure 6 depicts the chosen AST nodes. To convert the source code into AST, researchers use the Python package "java lang".

3.3 MAPPING TOKENS

Textual elements are generated for each file whenever the source code is translated. However, using them immediately as inputs into a DL framework is problematic. As a result, the text characters must be converted to an integers vector. The spectrum of one to the entire amount of characters is used to build the mapping between text characters and numbers. As a result, every word will be represented by a single number. To transform all textual characters into a series of numbers, the "Keras Text tokenization utility" module is used. The size of all integer vectors must be consistent. The length that has been determined would be chosen. As a result, if the vector size is less than the specified value, it is filled in with 0. Furthermore, if the vector size exceeds the specified value, the excess size will be eliminated.

3.4 CLASS IMBALANCE HANDLING

Majority of the cases, the information is unbalanced because the amount of damaged documents is lesser than the number of good documents. As a consequence, the forecasting outcomes would be labeled as clear because cleaned records make up most of the data. For handling class imbalance correlation clustering is performed here. Traditional ways of removing data imbalance include logistic regression, up-sampling, down-sampling, and over-sampling [17]. This approach is less accurate than PCA-based data imbalance elimination. PCA may be used to reduce a high-dimensional point to a lower-dimensional point, and then filters can be used to rank the relevance of the chosen features. The variance-covariance structure of a group of variables is described by principal component analysis (PCA) in terms of fewer new variables that are linear combinations of the original variables. The additional variables may be readily produced via eigenanalysis of the original data's covariance matrix or correlation matrix. It is advisable to do PCA on the sample correlation matrix if the variables are assessed on scales with widely differing ranges or if the units of measurement are not comparable. Finally, the PCA provides a new PC transform that is constructed by utilizing the data correlation matrix to select the best PCs among all of the features. Therefore, after encoding of data, PCA is applied to them to handle their class imbalance issue.

3.5 RELEVANT FEATURE EXTRACTION

The Data attributes are fed into this algorithm and relevant features are extracted. Finding the characteristics of features as well as selecting meaningful and non-redundant features from such information is a big challenge. Various evaluation metrics are used in feature selection features to find the optimal feature subset, yet they do not search for feature structures. Clustering is a more effective method that recognizes the structure of features and eliminates noisy or unnecessary information. In this paper, we have applied correlation clustering is used.

To uncover the structure of the database as well as build the clusters, the K-means clustering algorithm is utilized. As initial cluster centers, choose 'k' features at random from the original dataset D. The cluster is given to the most comparable item based on the distance between both the features and the cluster mean. Within every cluster, a new mean number is evaluated. The procedure was repeated till there was no further redistribution of features from any of the clusters. A user must define the number of clusters in advance when using the k-means clustering algorithm. In this work, two clusters are created on the whole dataset. The Euclidean distance function is used to determine the similarity between the two features, as indicated in Eq (1):

$$d_{a,b}{}^2 = (a_1 - b_1)^2 + (a_2 - b_2)^2$$
$$d_{a,b} = \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2} \quad (1)$$

Where, A= [a_1, a_2] and B= [b_1, b_2] are testing feature vectors.

After the clusters have been constructed, the features that do not fit into any of them are removed. The next step is to clear each cluster of duplicated features. This correlation filtering technique is employed to do this. Eq. (1) is used to compute the correlation between features. Pearson's linear correlation coefficient is defined for features P having values p as well as classes Q with values q, where P and Q are viewed as random variables:

$$\varrho(X, C) = \frac{E(PQ) - E(P)E(Q)}{\sqrt{\sigma^2}(P)\sqrt{\sigma^2}(Q)} \quad (2)$$

Algorithm 1: Correlation Clustering of Testing Metrics

Input: Data Attributes, $D\{d_1, d_2, d_3, \ldots d_n\}$
No. of clusters, n
Output: Optimal Features, F
Procedure
1: Begin
2: Initialize, k cluster centers
3: Calculate Euclidian Distance between D and k (Eq. 1)
4: Join closet cluster based on similarity
5: Compute new mean
6: Repeat steps 3 to 5 until the Convergence criteria met
7: Remove irrelevant features in D
8: For i:1 to i<=k
9: Calculate correlation (Eq. 2)
10: Select best features (F)
11: End
12: Return F
End

3.6 FINE-TUNED CNN CLASSIFICATION

In the fine-tuned model, the output from the last convolution block of ResNet-50 is cascaded with denseconv layers with parametric ReLU activation and batch normalization layers. This network is retrained and features were extracted from the proposed fine-tuned model and further flattened to classify the data. Fine-tuning was performed by applying a parametric ReLU activation layer and using a hybrid loss function. This loss function is designed to solve the problem of class imbalance. The loss function is described below:

$$Loss_H = \frac{1}{N}\sum_{i=1}^{I}\sum_{j=1}^{J} L_{ij} \quad (3)$$

Where, J = Number of loss functions, I = Number of layers in the model and L_{ij} = Focal loss

For fine-tuned model, parametric ReLU is adopted because it fine-tunes the learning parameters on its learning rate without any vanishing gradient problem.

4. RESULT ANALYSIS

4.1 DATASET DESCRIPTION

From the PROMISE database, seven open-source initiatives built in Java were selected. It's a publicly available database that's utilized to forecast software flaws. These initiatives include an XML converter, a textual web search engine, and data communication drivers, among others. Table 1 contains information on every initiative, such as its title, two releases, overall documents, and fault percentage. Every initiative has two deployments, the first of which is utilized for training the framework and the second of which is employed to evaluate it. The initiatives in PROMISE include standard characteristics for every Java file. Table 2 lists all of the testing metrics.

Table 1. Description of the PROMISE dataset

Project	Releases	Total files
camel	1.0, 1.2, 1.4, 1.6	2784
Jedit	3.2, 4.0, 4.1, 4.2, 4.3	1749
Poi	1.5, 2.0, 2.5, 3.0	1378
Synapse	1.1, 1.2	478
Xalan	2.4, 2.5, 2.6, 2.7	3320
Xerces	1.2, 1.3	893

Table 2. Testing Metrics of PROMISE Dataset

Testing Metrics	
Weighted methods per class	Lines of code
Depth of inheritance	Response of class
No. Of children	Data access metric
Coupling between object classes	Measure of aggregation
The measure of function abstraction	Afferent coupling
Lack of cohesion in methods	Efferent coupling
Cohesion among methods of a class	Inheritance coupling
Lack of cohesion in methods	Average method complexity
No, of public methods	Coupling between methods

4.2 EVALUATION METRICS

The confusion matrix is utilized to demonstrate the performances of proposed frameworks. It incorporates the forecasting findings of the framework. It also produces the following groups of outcomes: True Positive (TP), False Positive (FP), True Negative (TN), and False Negative (FN) (FN). Where TP represents the number of anticipated faulty codes which are already faulty, TN represents the number of anticipated non-faulty codes which are already non-faulty, FP represents the amount of anticipated faulty documents that are correct, and FN represents the amount of anticipated correct documents that are faulty. The used Metrics in this paper are discussed below:

$$\text{Accuracy} = (\text{TP} + \text{TN})/\text{TP} + \text{TN} + \text{FP} + \text{FN} \quad (4)$$

$$\text{Recall} = \text{TP}/\ (\text{TP}+\text{FN}) \tag{5}$$

$$\text{Precision} = \text{TP}/\ (\text{TP}+\text{FP}) \tag{6}$$

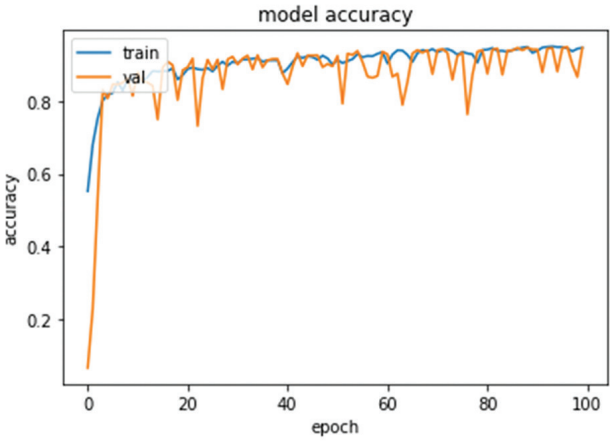
$$\text{F1\_Score} = \frac{2 * (\text{Precision} * \text{Recall})}{(\text{Precision} + \text{Recall})} \tag{7}$$

4. 3 IMPLEMENTATION DETAILS

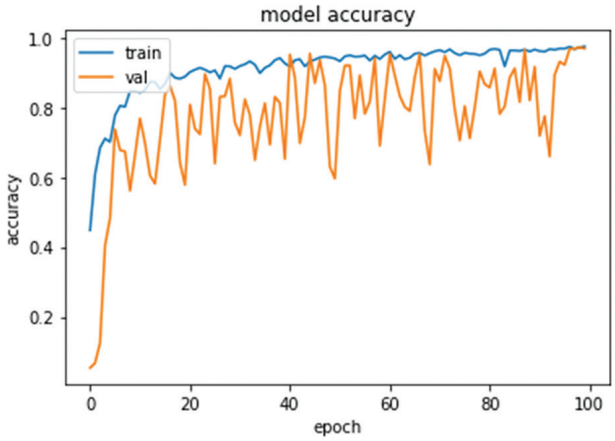
This section presents the training details of the proposed software testing model for fault prediction. For the experiment, we used python for training and testing purposes. For this, the entire dataset is divided into a 70:30 ratio. The minimum batch size for the training autoencoder was taken to be 128. The maximum epoch was taken to be 100. Notably, these model utilizes approx. 10GB RAM on high-performance GPU computing. Therefore, this model was trained on computing service provided by google i.e., google colab.

4. 4 RESULT ANALYSIS

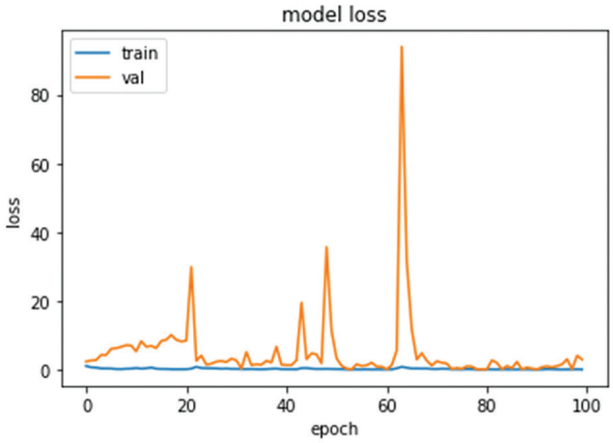
Fig-3(a)-3(f) shows the model accuracy for six datasets Camel Dataset, Jedit dataset, Poi dataset, Synapse dataset, Xerces dataset, and Xalan dataset. The accuracy is given for training and validation sets. The epoch varied from 0-100 at the x-axis. Because for every graph the convergence rate is for less than 100 epochs. Fig-3(a) shows the training and validation accuracy for Camel Dataset. The graph converges at initial epochs nearly in between 20-40. The average accuracy for Camel dataset is 95 %. Fig-3(b) shows the training vs validation accuracy over the number of epochs varied from 0-100 for JEDIT dataset. The accuracy varied from 0-20 epochs and then start to converge after 20 epochs. The average accuracy for JEDIT dataset is 97%. For the Synapse dataset, Fig-3(c) demonstrates the training versus validation accuracy across a range of epochs from 0-100. The accuracy ranged from 0 to 60 epochs, and after 60 epochs, it began to converge. The Synapse dataset has an average accuracy of 88 percent.Fig-3(d) shows the training vs validation accuracy for the Poi dataset over a range of epochs from 0-100. The accuracy ranged from 0 to 20 epochs, and it started to converge after 20 epochs. The Poi dataset has a 79 percent accuracy rate. The training vs validation accuracy over a range of epochs from 0-100 is shown in Fig-3(e) for the Xalan dataset. The precision fluctuated from 0 to 20 epochs, and it started to converge before 20 epochs. The accuracy of the Xalan dataset is 77 percent on average. Fig-3(f) shows the training vs validation accuracy for the Xerces dataset across a range of epochs from 0-100. The accuracy fluctuated from 0 to 20 epochs, with the accuracy beginning to converge after 20 epochs. The Xerces dataset is 95.4 percent accurate on average.



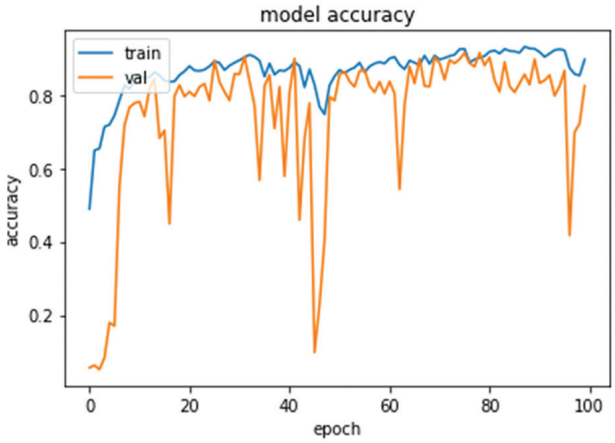
(a)



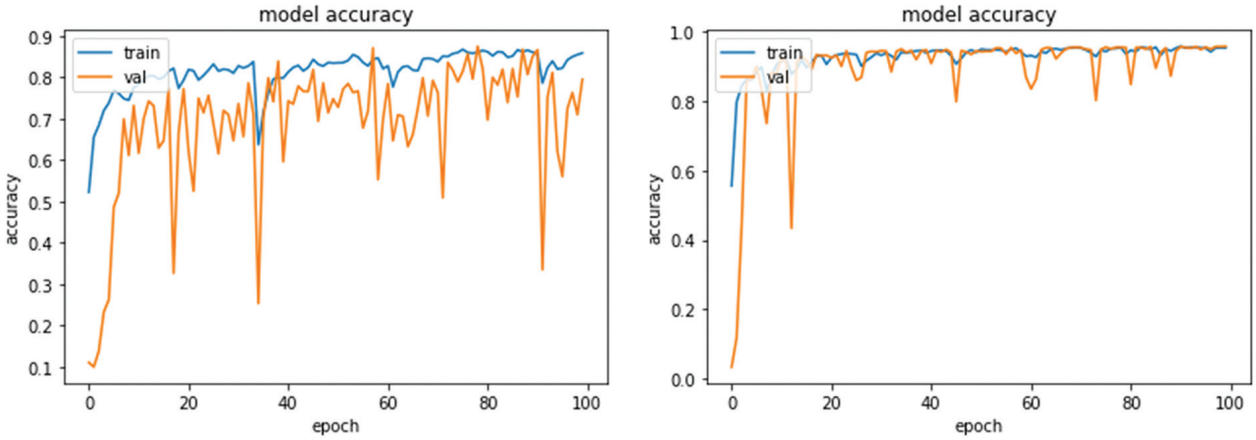
(b)



(c)



(d)



(e) (f)

Fig. 3. Training and Validation Accuracy over Given Datasets

Table 3 shows the Accuracy, Precision, Recall, and F–Measure comparison of all the datasets. For camel dataset Accuracy, Precision, Recall, and F–Measure is 95 %. For Jedit dataset accuracy is 97 %, precision is 96.7 %, Recall is 96.9% and F-Measure is 97 %. For Poi dataset accuracy is 82 %, precision is 80 %, Recall is 82 % and F-Measure is 79 %. For Synapse dataset accuracy is 82 %, precision is 77 %, Recall is 81% and F-Measure is 78 %. For Xalan dataset accuracy is 80 %, precision and Recall are 79 % and F-Measure is 77 %. For Xerces dataset accuracy is 96 %, precision is 95.6 %, Recall is 96 % and F-Measure is 95.4 %. Overall for all dataset average accuracy is 88.6 %, average precision is 85.5 %, average Recall is 88.3 % and average F-Measure is 86.9 %. The maximum accuracy I attained with the Jedit dataset i.e. 97 %. Maximum Precision is attained by Jedit i.e.96.7 %. Maximum recall and F-measure are also maximum for Jedit Dataset around 96.9 % and 97 % respectively. Table 4 shows the comparison of different techniques discussed in [15] like RF, DBN, CNN, RNN, and BCIL along with our proposed work. It is clear from the table that our proposed method surpasses all the techniques given in [15] for all six datasets used in our study. For camel dataset, our evaluated f-measure is 95 %, which is 55.4 % higher than the lowest F-measure attained in RF[15]. Similarly, for Jedit dataset F-measure is 97% which is 49% higher than DBN[15] technique. For poi dataset, our proposed work is 13% better than the RF dataset. For Xalan F-measure is 77% and for Xerces it is 95.4 which is 77% better than the RF[15].

Table 3. Performance of Proposed Model

	Accuracy	Precision	Recall	F_measure
Camel	0.95	0.95	0.95	0.95
Jedit	0.97	0.967	0.969	0.97
Poi	0.82	0.80	0.82	0.79
Synapse	0.82	0.77	0.81	0.78
Xalan	0.80	0.79	0.79	0.77
Xerces	0.96	0.95.6	0.96	0.954
Average	0.886	0.855	0.883	0.869

Table 4. F-measure of Proposed and State-of-art models

Project	RF [15]	DBN [15]	CNN [15]	RNN [15]	CBIL [15]	SMO [18]	Ours
Camel	0.396	0.335	0.505	0.515	0.935	0.427	0.95
Jedit	0.550	0.480	0.631	0.595	0.850	0.734	0.97
Poi	0.669	0.780	0.778	0.722	0.852	0.390	0.79
Synapse	0.414	0.503	0.512	0.487	0.889	0.617	0.88
Xalan	0.638	0.681	0.676	0.606	0.716	0.432	0.77
Xerces	0.185	0.261	0.311	0.262	0.951	0.612	0.954
Average	0.475	0.506	0.5688	0.5311	0.865	0.535	0.885

5. CONCLUSION

Software testing is the most critical element in the development life cycle since software systems are continually developing. Deep learning has subsequently made significant progress in automatically extracting semantic characteristics and improving software defects predictions performance and accuracy. This study presents the Correlation Clustering fine-tuned CNN (CCFT-CNN) model. It aids in improving code review and software testing by predicting the parts of source code that are faulty. The CCFT-CNN model is tested on the Camel Dataset, Jedit Dataset, Poi Dataset, Synapse Dataset, Xerces Dataset, and Xalan Dataset from the PROMISE dataset. The findings show that CCFT-CNN outperforms the existing models of Random Forest, DBN, CNN, RNN, and CBIL in terms of performance. The proposed CCFT-CNN outperforms the RF model, which has the lowest F-Measure in the baseline model, by 41%. For the DBN, CNN, and RNN models, the CCFT-CNN improves by 37.9%, 34.7 percent, and 35.4 percent, respectively. CBIL, on the other hand, has the best performance in the baseline model, and its comparison improves the outcome by 2%. Quality measures such as defect density, defect kinds, and defect severity should be included in the future. CCFT-CNN may also be used on a variety of open-source projects developed in a va-

riety of programming languages. Furthermore, to improve the quality of public datasets, new data preparation methods should be included.

6. REFERENCES

[1] F. Salfner, M. Lenk, M. Malek, "A survey of online failure prediction methods", ACM Computing Surveys, Vol. 42, No. 3, 2010.

[2] N. S. Chauhan and A. Saxena, "A green software development life cycle for cloud computing", IT Professional, Vol. 15, No. 1, 2013, pp. 28-34.

[3] J. Wang et al. "Robot: robustness-oriented testing for deep learning systems", Proceedings of the IEEE/ACM 43rd International Conference on Software Engineering, Madrid, Spain, 22-30 May 2021.

[4] J. W. Duran and S. C. Ntafos, "An evaluation of random testing," IEEE Transactions on Software Engineering, Vol. SE-10, No. 4, 1984, pp. 438-444.

[5] S. Anand et al. "An orchestrated survey of methodologies for automated software test case generation", Journal of Systems and Software, Vol. 86, No. 8, 2013, pp. 1978-2001.

[6] H. B. Braiek, F. Khomh, "On testing machine learning programs", arXiv:1812.02257, 2018.

[7] S. Abrecht et al. "Testing deep learning-based visual perception for automated driving", ACM Transactions on Cyber-Physical Systems, Vol. 5, No. 4, 2021, pp. 1-28.

[8] S. Gerasimou et al. "Importance-driven deep learning system testing", Proceedings of the IEEE/ACM 42nd International Conference on Software Engineering, Seoul, Korea, 5-11 October 2020.

[9] J. Wang, et al. "Robot: robustness-oriented testing for deep learning systems Proceedings of the IEEE/ACM 43rd International Conference on Software Engineering, Madrid, Spain, 22-30 May 2021.

[10] Q. Guo et al. "Audee: Automated testing for deep learning frameworks", Proceedings of the 35th IEEE/ACM International Conference on Automated Software Engineering, Melbourne, VIC, Australia, 21-25 September 2020.

[11] L. Ma et al. "Deepmutation: Mutation testing of deep learning systems", Proceedings of the IEEE 29th International Symposium on Software Reliability Engineering, Memphis, TN, USA, 15-18 October 2018.

[12] L. Ma et al. "Deepgauge: Multi-granularity testing criteria for deep learning systems", Proceedings of the 33rd ACM/IEEE International Conference on Automated Software Engineering, September 2018, pp. 120-131.

[13] X. Du et al. "Deepcruiser: Automated guided testing for stateful deep learning systems", arXiv:1812.05339, 2018.

[14] J. Wang et al. "Adversarial sample detection for deep neural network through model mutation testing", Proceedings of the IEEE/ACM 41st International Conference on Software Engineering, Montreal, QC, Canada, 25-31 May 2019.

[15] A. B. Farid, E. M. Fathy, A. S. Eldin, L. A. Abd-Elmegid, "Software defect prediction using hybrid model (CBIL) of convolutional neural network (CNN) and bidirectional long short-term memory (Bi-LSTM)", PeerJ Computer Science, Vol. 7, 2021, pp. 1-22.

[16] Y. Fan, X. Xia, D. Lo, A. E. Hassan, Y. Wang, S. Li, "A Differential Testing Approach for Evaluating Abstract Syntax Tree Mapping Algorithms", Proceedings of the IEEE/ACM 43rd International Conference on Software Engineering, Madrid, Spain, 22-30 May 2021., pp. 1174-1185.

[17] S. Goyal, "Handling Class-Imbalance with KNN (Neighbourhood) Under-Sampling for Software Defect Prediction", Artificial Intelligence Review 2021 553, Vol. 55, No. 3, 2021, pp. 2023-2064.

[18] Z. Sun, J. Zhang, H. Sun, X. Zhu, "Collaborative filtering based recommendation of sampling methods for software defect prediction", Applied Soft Computing, Vol. 90, 2020, p. 106163.

Information retrieval approaches: A comparative study

Case Study

Assmaa Moutaoukkil
IRF-SIC Laboratory, Department of Computer Science, Faculty of science,
Ibn Zohr University, Agadir, Morocco
Asmae.mt1@gmail.com

Ali Idarrou
IRF-SIC Laboratory, Department of Computer Science, Faculty of Science,
Ibn Zohr University, Agadir, Morocco
aliidarrou@gmail.com

Imane Belahyane
IRF-SIC Laboratory, Department of Computer Science, Faculty of Science,
Ibn Zohr University, Agadir, Morocco
imane.belahyane@edu.ac.uiz.ma

Abstract *– The area of information retrieval (IR) has taken on increasing importance in recent years. This field is now of interest to large communities in several application domains (security, medicine, aeronautics, etc.). IR studies find relevant information from the semi-unstructured type of data. As the information resources generated after the search can be extensive in quantity and different in quality, it is essential to rank these results according to the degree of relevance. This paper focuses on text information retrieval (TIR) and emphasizes the importance of each IR approach. This study presents insightful aspects of TIR and provides a comparative study between some proposed approaches and models. Each model offers IR advantages and suffers from several limitations.*

Keywords*: information retrieval, NLP, graph, TF-IDF, document classification, KNN, fuzzy K-means, deep learning, BERT, transformers,*

1. INTRODUCTION

The increasing growth of technologies, storage capacity and computational power, and modification of telecommunication technologies provided to deal with different types of information maintained by different types of media. The amount of online data has grown at least as quickly as the speed of computers, and we would now like to be able to search collections that total in the order of billions to trillions of words. Consequently, several procedures and methods are needed to find the relevant information at the right time.

Information has always been the raw material of all processing for the benefit of organizations and individuals. It is created, processed, shared, stored, and transmitted by these units loudly. The continuity and development of any legal or physical person or unit are linked to its ability to access relevant information at the right time. While the digital revolution has undoubtedly reduced the importance of time and distance, in some ways, it has grown it in others: IR and managing large corpora and collections have become significant challenges. Many applications that deal with information would be inadequate without the support of IR technology. IR had occupied information scientists before the term "information science" was coined [1]. Similarly, information science, such as text mining, human-machine interaction, NLP, machine learning, Big Data, etc., is in the service of IR. Moreover, if the information is an integral part of today's world, IR is a decisive activity for the people who inhabit it. The tasks related to this domain seek to automatically respond to the information needs of a user who hopes to solve a problem or achieve a goal for which the current state of his knowledge is inadequate [2]. If the idea is simple, the methods are much more delicate, especially when dealing with complex data, for example, by their modality, as is the case of the textual contents which constitute the object of study of this work. One way to clarify the problems treated in this work is to ask these questions:

- How to capture and model the need initially expressed by the user to fully satisfy it?
- How can content be optimally represented to facilitate the IR process?
- How to access the relevant documents requested by the user through a query in an IR system?

- How to compare and class documents in a dynamic, rich, unstructured, and voluminous collection?

There is no single convincing answer to these questions. There are many approaches, called here models, and each is useful for developing some IR tools. In sections 4,5,6, and 7, we will explain in a pedagogical style, respectively, the models based on bag-of-words representation, those based on graphs, those based on classification techniques, and those based on deep learning techniques. We will devote section 8 to discussing and comparing these models. We explain in section 3 the methodology followed for the elaboration of this work. But first, we will describe what exactly these models denote (section 2).

2. INFORMATION RETRIEVAL(IR)

IR refers to the process, methods, and procedures for locating and extracting data and information stored in a semi-structured or unstructured database.

Gerard Salton defined IR as follows: "Information retrieval is a field concerned with the structure, analysis, organization, storage, searching, and retrieval of information." [3]

IR is the science of searching for information in a document or documents themselves and searching for the metadata that describes the data and databases of text, images, or sounds.IR is the science of obtaining information system resources relevant to an information need from a collection of such resources. In this paper, we deal with textual data only.

2.1. IR PROCESS

Information Retrieval Systems (IRS) are concerned with search strategies in which the retrieved documents may be more or less relevant to the query. There are four basic processes that an IRS must support: *indexing*, *interrogation*, *similarity evaluation*, and *ranking*. These processes are visualized in Fig. 1, where the square boxes represent objects, and the rounded boxes represent processes.

The process of IR begins with the representation of information. The elementary object of information storage is the document. The representation of documents is generally referred to as the indexing process. This process takes place offline, i.e., the end-user of the IRS is not directly involved. The indexing process results in a representation of the document in the first step (1) the represented documents are then stored (2) the query, in turn, must be represented in the same way as the documents (5)(6) The IRS then proceeds to a comparison process through a similarity measure which changes depending on the IR model used (7)(8). The ranking process aims at sorting the documents deemed relevant by the system (9). The user can then judge the relevance of the returned documents based on either his initial information need or a new information need. An interrogation process assists the user in making his vague and ambiguous need explicit in the form of a query (3).

In order to optimize the results of an IRS, a query expansion process is always recommended, either by explicit or implicit feedback.

3. METHODOLOGY

This work was conducted following the guidelines suggested by [4][2][5]. The process followed is illustrated in Fig. 2. It should be noted that it is sometimes difficult to compare the results published in different papers due to incomplete or insufficient information about the algorithm used by the authors, e.g., the pre-processing methods adopted may lead to significant differences. Therefore, whenever possible, we collect results from papers that contain comparisons between some of these models performed on a single site for reliability reasons. We compare the empirical evaluation results of the most prominent IR models previously discussed on several popular reference datasets. Nevertheless, some previous studies are not based on popular corpora, and in this case, our analysis is based on other additional criteria, namely:

- The nature, number, and category of the documents in the collection.
- We have also specified criteria for selecting papers to facilitate comparing these models (comparison of the proposed contribution with previous models, ...)

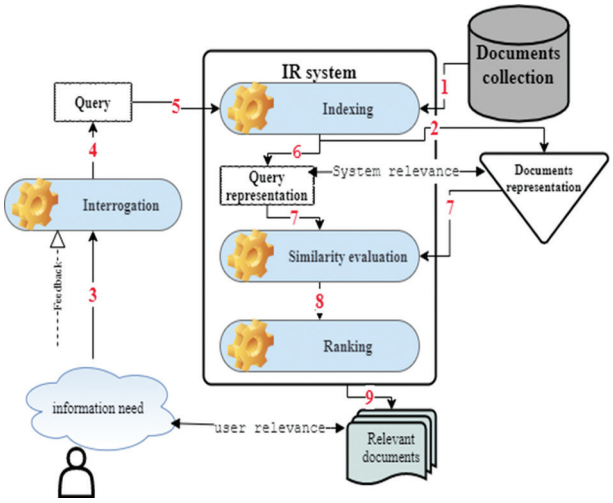


Fig. 1. IR process

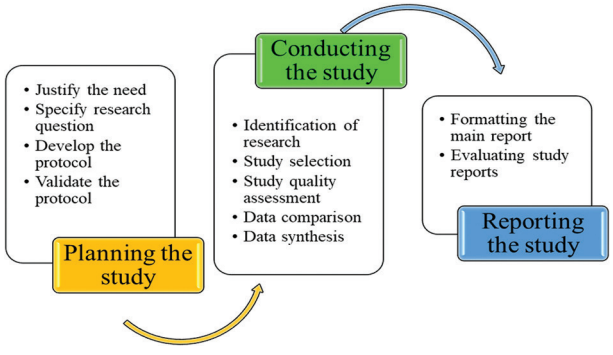


Fig. 2. Study process

4. MODELS BASED ON THE BAG-OF-WORDS REPRESENTATION

What is important to users is the content of a document. Given that the content is semi-structured or unstructured, the challenge for an IRS is to produce a set of words or terms that are sufficiently descriptive to represent the content of a document, in addition, to measuring the importance of each word in a particular document and the collection in general.

4.1 BAG OF WORDS REPRESENTATION

The initial idea is to use keywords to identify the content of a document[6] as well as the content of the query. The process starts by listing all the significant words in each document to identify the keywords in the collection. The goal is to find the words that best describe the content of a document.

Bag-of-words representation provides a way to organize the textual content of a collection of documents into a matrix (Table 1) in the hope of mapping key terms (words) to documents to measure the importance of each term in each document in the collection using weighting techniques.

Table 1. bag of words representation

	d1	d2	....	dn-1	dn
t1	$W_{1,1}$	$W_{1,2}$	.....	$W_{1,n-1}$	$W_{1,n}$
t2	$W_{2,1}$	$W_{2,2}$	...	$W_{2,n-1}$	$W_{2,n}$
⋮ ⋮	⋮ ⋮ ⋮	⋮ ⋮ ⋮	⋮	⋮ ⋮ ⋮	· ⋮ ⋮ ⋮
tn-1	$W_{n-1,1}$	$W_{n-1,2}$	....	$W_{n-1,n-1}$	$W_{n-1,n}$
tn	$W_{n,1}$	$W_{n,2}$	....	$W_{n,n-1}$	$W_{n,n}$

In order to assign a weight to each term, many weighting solutions have been proposed in the literature [7].The most impressive is the TF-IDF. Furthermore, since a demonstration is better than a long explanation, we present the following example:

D {d1,d2,.....,dn} a collection of documents.

T {t1,t2,.....,tn } of terms belonging to this collection without redundancy.

$W_{t,d}$ The statistical measure that reflects the importance of a term t in document d.

4.2 WHY THE TF-IDF METHOD?

TF-IDF is a weighting method that measures the importance of a word in a document relative to a collection of documents. It is a statistical measure that is based on the calculation of the weight of each term in a document relative to a collection of documents according to the following general formula:

$$W(t,d) = TF \times IDF \quad (1)$$

TF (Term frequency) is the frequency of appearance of a term t in document d. The number of times that t appears in d. many variants have been proposed in the literature. The simplest is the Boolean model; 0 if the term is present in document 1 otherwise, this approach is very limited. Others propose a raw frequency or a logarithmic normalization to dampen the differences or a normalization that considers the length of the document.

The IDF has a discriminating power. IDF (inverse document frequency) measures how common or rare a term is in a collection of document D. It allows assigning a high weight to less frequent terms and a lower weight to more frequent terms. Fig. 3 and 4 show the results of an experiment that proved the usefulness of using both TF and IDF [8]. The collection studied was 1400 documents encoded in SGML text format, keeping the formatting tags to account for noisy data and to test the robustness of TF-IDF. The authors of [8] apply case sensitivity to simulate more noise. The proposed SRI returns the first 100 documents as follows:

It is quite clear that TF-IDF is more powerful than TF alone. It provides high values for rare words and low values for common words. TF returns imprecise results; relevant documents are scattered sporadically. This can be explained by IDF eliminating the more frequent stop-words.

The results in [9] show that the TF-IDF system managed to have an accuracy that exceeds that achieved by the TF-ATO system in the first case without a discriminative approach or stop-word removal. This is explained by the fact that the TF-IDF system can remove insignificant terms (IDF).

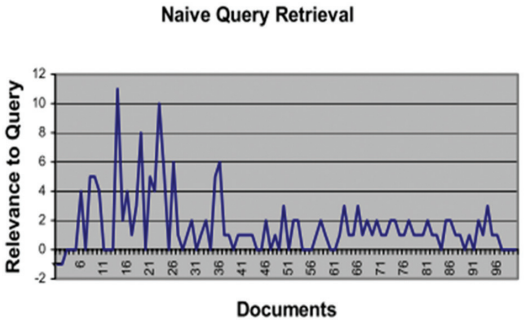


Fig. 3. Result using TF alone to calculate the weights [8]

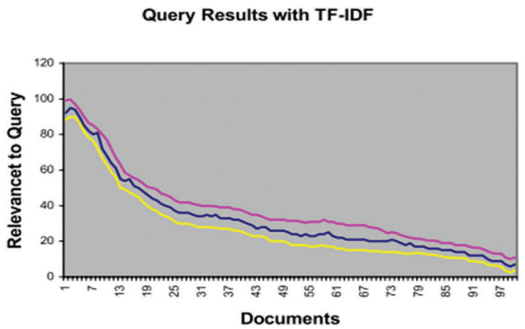


Fig. 4. Result using TF-IDF to calculate weights [8]

4.3 PREPROCESSING DATA TO IMPROVE PERFORMANCE

In written language, some terms carry more semantics than others. Several authors claim that there is still interest in improving the simple TF-IDF. Looking at the data, the simplest way to improve TF-IDF would be to ignore case sensitivity [8] and to use techniques to clean and normalize the data, such as Stemming: which is used to group the different forms of a particular word as well as to assimilate the words with their lexical derivatives and synonyms.

To limit the amount of useful information and the term space, several authors propose a preprocessing process to eliminate all non-significant words or reduce any other noise (upper case, lower case, suffix, prefix). As an example, the authors of [10] propose to remove stop-words such as "and", "a", and "an" because they are non-significant words, not having a discriminating character. They do not make it possible to distinguish one text from another.

Indexing is a crucial step in the IR process. In a keyword-based approach, this technique allows representing a document and/or a query by a set of keywords, also called descriptors. These descriptors constitute an exploitation facility of the documents by the IRS. Given the large amount of data to be processed by this system, preprocessing techniques can reduce the limits of this approach by cleaning the data, such as eliminating stop words, the stemming of Porter, Tokenization, etc. Nevertheless, preprocessing is not sufficient to increase the performance of an IRS.

5. GRAPH-BASED MODELS

A standard approach to IR is to model the text as a bag of words. Alternatively, the text can be modeled as a graph, whose vertices represent words and edges represent relationships between words, defined based on any meaningful criterion.

With such a graph, graph-theoretic calculations can be applied to measure various properties of the text.

The proposed idea is to transform IR into a graph problem. The graph data structure allows organizing and especially linking a set of objects (documents, terms, queries...) simply and practically.

A variety of models and techniques have been proposed for this purpose:

In [11] the authors present the graph-based TextRank model for extracting keywords and phrases from raw data.

The algorithm considers the weight of edges when computing the score associated with a graph vertex.

In [12] the authors introduce the use of a random walk algorithm to weight the terms in the TF-IDF weighting scheme by adapting the Text Rank algorithm. In this model, the edge weights are not considered.

In [13] the authors exploit the relationship between the local information of a vertex (term position) and the global information (information gain) and term dependence to produce term weights.

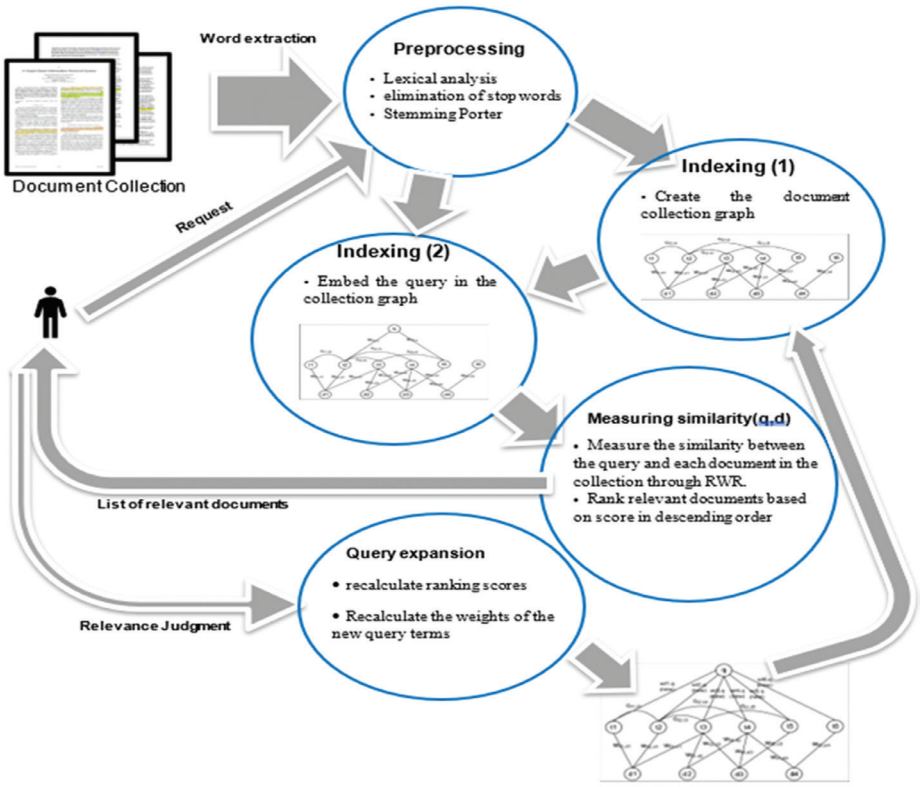


Fig. 5. IR graph-based model

For example, we will look at the model proposed in [14]. This model adopts a graph structure that captures the occurrences of terms in documents and the correlations between terms. The similarity between a document and the query is computed by a specific type of random walk graph algorithm called Random Walk with Restart (RWR). The model can be extended to incorporate relevant feedback naturally. Its effectiveness is evaluated using TREC collections; its performance is measured and compared to systems participating in the TREC program. The diagram in Fig.5 summarizes the proposed process.

The model builds a collection graph that represents the content of documents and the relationships between terms created before a query exists. When a user provides a query to the system, a query node is constructed and connected to the query terms in the graph. A graph random walk algorithm is then applied to calculate the relevance of each document to the query. The present model was tested on the TREC 2003 High Accuracy Retrieval from Documents (HARD) corpus (National Institute of Standards and Technology, 2009), which provides 20 training topics with 100 documents judged for each topic and uses a set of 48 topics for evaluation. The soft evaluation judgment is adopted in this experiment since the author does not use any additional information provided.

Four GIR configurations are studied:

The GIR-NoRelNoSig (Baseline) configuration uses binary link weights and does not include a relationship between terms.

The GIR-RelNoSig (corrBinary) configuration includes relationships between terms with binary link weights.

The GIR-RelSig (Full) configuration includes a term relationship with real-valued connection weights between 0.0 and 1.0.

The GIR-RF configuration is GIR-RelSig (RF) with the automatic relevance return process. The initial run's first ten documents are automatically fed into the model.

Comparing the four configurations of the GIR model, the performance was better from the weakest to the strongest GIR-NoRelNoSig, GIR-RelNoSig, and GIR-RelSig, respectively.

6. CLASSIFICATION TECHNIQUES BASED MODELS

The idea here is to transform the IR problem into a classification problem. As shown below in Fig. 6, the query is considered a document; we have a set of determined classes. Once a new document arrives, we must put this document in the suitable class or classes.

Dynamic document classification reduces effort and time as it processes the new document and assigns it directly to the appropriate classes without re-running the entire algorithm. [15]

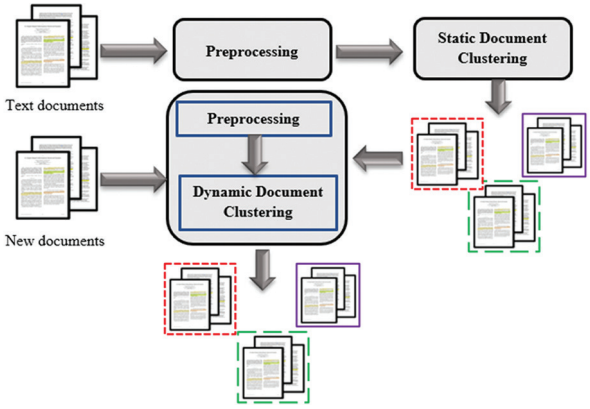


Fig. 6. A general process of dynamic document classification

Classification as a learning method automatically assigns one or more classes to a free text document. Document classification can be performed in supervised or unsupervised learning mode: "Supervised classification aims to associate each of the n observations {x1, ..., xn} with one of the k classes known a priori while unsupervised classification aims to group these data into k homogeneous groups" [16]

6.1 MODEL-BASED ON SUPERVISED CLASSIFICATION

Supervised classification is based on input and output examples to classify new documents. Predefined classes are assigned to text documents based on their content. The typical text classification process consists of *preprocessing*, *indexing*, *dimension reduction*, and *classification* [17]. The goal of classification is to train a classifier based on known examples, and then unknown examples are assigned automatically. Statistical classification techniques such as Naïve Bayesian classifier, nearest neighbor classifier, decision tree, and support vector machines can be used to classify the text. In [18] authors proposed a combination of TF-IDF and the KNN algorithm for classification. The optimized method performed six times better than the conventional method. Nevertheless, the quality of classification decreases slightly with the increase in the number of documents. The studied collection is 500 documents of different lengths and different categories. The following table (Table 2) summarizes the implementation of this algorithm and the result of this classification:

Table 2 . KNN and TF-IDF

Representation	**Bag of words**			
Preprocessing	⇨ adjust documents to classic text format ⇨ automatically remove control characters			
TF-IDF Formula	$a_{ij} = tf_{ij}idf_i = \frac{f_{ij}}{\sqrt{\sum_{s=1}^{N}\left(tfidf(a_{sj})\right)^2}} \times log_2\left(\frac{N}{df_i}\right)$			
Measure of similarity	**Euclidean distance**			
Documents categories	Sport	Politics	Finance	Daily news
Classification quality	0,92%	0,90%	0,78%	0,65%

This algorithm presents accurate results for the category "*sport*" the main reason for the excellent classification results in this category is that the documents were not textually demanding. The documents did not contain many different words, which reduced the impact of unusable words and characters. On the other hand, the worst classification was for the "*Daily News*" category. The content analysis of the documents in this category showed that these documents contained a lot of *unusable words*, those words that are often repeated and do not have a significant weight but harm the classification.

6.2 MODEL-BASED ON UNSUPERVISED CLASSIFICATION

Unsupervised classification or clustering is used to group similar documents. This method is based on the concept of dividing a similar text into the same cluster. Each cluster contains many similar documents [19] . The objects are grouped or clustered based on maximizing intra-class similarity and minimizing inter-class similarity.[20]

The authors of [15] proposed a comparative experiment between the two clustering techniques, HAC (hierarchical agglomerative clustering) and fuzzy K-means, for the same collection of 45 documents of different categories. The following table (Table 3) summarizes the algorithm followed along and the results obtained.

Table 3. HAC Vs fuzzy K-means

Representation		Bag of words			
Preprocessing		⇨ Eliminate StopWords ⇨ Stemming			
TF-IDF Formula		Nothing to report			
Measure of similarity		Cosine distance			
Documents categories		News 20	Reuters	Research papers	E-mail
HAC	Entropy	0.256213	0.112368	0.678951	0.225641
	F Measure	0.8612 31	0.8890 23	0.5234 10	0.7456 12
Fuzzy K-means	Entropy	0.389763	0.421189	0.245612	0.214561
	F Measure	0.785621	0.614523	0.884532	0.754312
Number of Clusters		4	6	4	4

It appears that HAC *New 20* and *Reuters data* perform better. On the other hand, Fuzzy K-means is more accurate for the Research Paper data. Either algorithm can be applied to the E-mail dataset, as the entropy and accuracy values do not show a significant difference. However, the graph (Fig. 7) indicates that e-mails can be well classified with Fuzzy k-means, which produces non-overlapping clusters.

Thus, the impact of document categories on the performance of either supervised or unsupervised classification is infallibly proven.

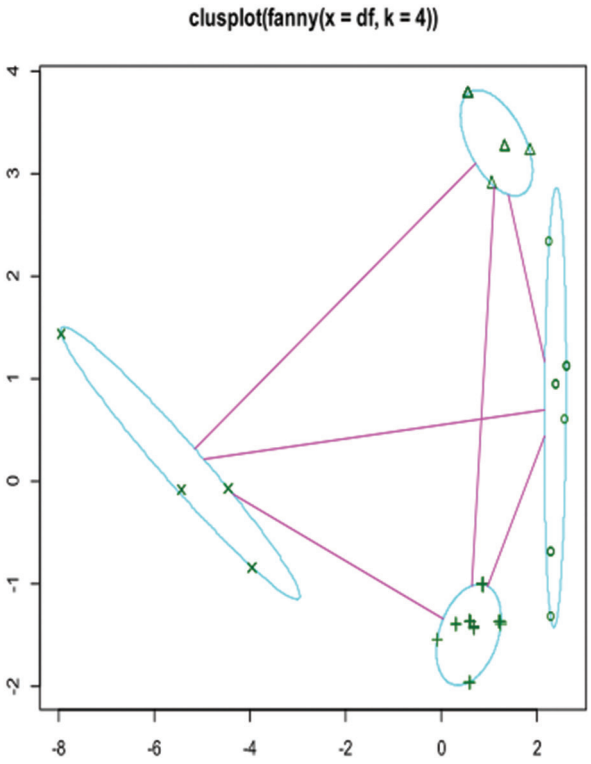


Fig. 7. Fuzzy K-means plots for email dataset[15]

6. 3 CHOOSING THE RIGHT SIMILARITY MEASURE

An experiment in [21] aims to find the best distance measure of hierarchical clustering techniques (HAC), K-means, and K-medoids. Many distance measures were tested: 'Euclidean distance', 'Cosine distance', 'Distance between neighborhoods', 'Hamming distance', 'Jaccard distance', 'Correlation distance', 'Chebyshev distance' and 'Sqeuclidean distance'.

The collection comprises five categories: Business, Education, Elections, Entertainment, and Games. Each category has 50 articles selected for implementation, which are taken from the websites of different news channels. In hierarchical clustering, Euclidean distance measure by means was the best. In K-means clustering, correlation distance measure, and K-medoids clustering, city block distance measure shows better results than other distance measures.

The HAC evaluation results show that the *classification quality using Euclidean distance exceeds that of Cosine distance by 0.2597*. Hence, the usefulness of questioning the results of using Euclidean similarity measure in the HAC technique of the algorithm proposed by [15].

7. DEEP LEARNING BASED MODELS

For text classification, after classification learning came deep learning. In the context of IR, deep learning approaches were attractive for two reasons: First, continuous vector representations freed text retrieval from the limitations of exact term matching. Second,

neural networks promised to avoid the need for laboriously hand-crafted features (which addressed a major difficulty in building systems using rank-based learning)[22] . Deep learning models do not take raw text as input; they only work with numerical tensors by vectorizing text into numerical tensors. [23]

In the literature, several models of neural networks applied to IR are used in combination to achieve overall efficiency. Table 4 presents an overview of the results obtained. The citation in each row indicates the original paper where the method is proposed. The backslash symbols indicate no published results for the specific model on the specific dataset in the related literature. Based on the reported results, we observe that:

Table 4. Overview of previously published results on ad-hoc IR datasets.

	Robuste04		ClueWeb-09-Cat-B	
	CARD	P@20	CARD	P@20
BM25 [24]	0.255	0.370	0.101	0.326
QL [25]	0.253	0.369	0.100	0.328
DSSM [26]	0.095	0.171	0.039	0.131
CDSSM [27]	0.067	0.125	0.054	0.177
ARC-I [28]	0.041	0.065	0.024	0.089
ARC-II [28]	0.067	0.128	0.033	0.123
MP [29]	0.189	0.290	0.066	0.158
DRMM [25]	0.279	0.382	0.113	0.365
PACRR [30]	0.254	0.363	/	/
NPRF-KNRM [31]	0.285	0.393	/	/
NPRF-DRMM [31]	0.290	0.406	/	/
BERTBase MaxP[22]	0.365	0.465	/	/
BERTLarge MaxP[22]	0.374	0.477	/	/
BERT-QE-Large[22]	0.386	0.489	/	/
BERT-QE-Medium[22]	0.383	0.489	/	/
PARADE[22]	0.380	/	/	/

- Although the probabilistic models (QL and BM25) are simple. They can already achieve reasonably good performance.
- The asymmetric, interaction-oriented, multi-granularity architecture can perform better than the symmetric, representation-oriented, single-granularity architecture on ad-hoc search tasks, except for SNRM.
- As a specific instance of transformer architectures, BERT provides superior results to what has gone before. It is a robust empirical result that is widely replicated. BERT stands out for bringing together many crucial ingredients to produce tremendous advances in inefficiency. As a more sophisticated model, it draws many vital insights: the goal of contextual integrations is to capture complex features of language (e.g., syntax and semantics), as well as the way meanings vary across linguistic contexts (e.g., polysemy).

8. DISCUSSION

The IR models analyzed in this work have several contributions and suffer from several limitations:

A model based on the bag-of-words representation is simple, efficient, and easy to implement, which makes it ideal for forming the basis of more complex algorithms and IRS [32]. However, the bag-of-words representation and TF-IDF are constrained by several challenges:

- The TF-IDF treats words according to their morphology. For example, “year” and “years” will be considered as two separate words, which leads to a decrease in the weight of the word in the collection.
- An ample term space slows down the search and consumes memory space.
- The effectiveness of TF-IDF decreases with increasing collection scale[8][33]

Based on a bag-of-words representation, this approach does not consider the relationship between words. A word can have several synonyms [34]. Moreover, TF-IDF is limited to a lexical function and does not allow checking text semantics. So, it is not practical to search for co-occurrences of a word [6][8].

A graph-based model has the advantage of linking multiple and various objects with various fast and scalable techniques. Its structure is flexible to incorporate many performance improvement techniques. The results show that including term relationships and term importance weighting is helpful for search. However, incorporating automatic query expansion into the model is not very useful. Nevertheless, with the crazy evolution of the documents to be indexed before processing and the dynamic nature of the collection, the indexing space becomes larger and larger. In addition, the high frequency of use requires asking questions about the response time, the adaptability of a model, and the availability of relevant information at the right time. We mean by the adaptability of a model its ability to progressively update the indexing of the database with the constant operation when the database is periodically updated with new information[35].

A model based on classification techniques can remedy its limitations and has several advantages:

- Accelerated search process:
  number of classes < number of documents
- Adaptability and extensibility for dynamic collections to find relevant information in response to an information need, during a mass of constantly changing data.

On the other hand, the techniques treated in the present synthesis are based on statistical approaches using TF-IDF. It is then limited to a lexical function of words without considering the problems related to ambiguities. Hence the importance of observing the

problem of the semantic gap and the possibilities of enriching these models with solutions that allow for the semantic aspect.

When faced with such a problem to study the meaning of a word ideally, it should be observed in its context. Thus, in a text (and by extension in language), there is more or less significant dependence between words.

A recent family of techniques (circa 2013) has rethought models with a representation of words in a space with some similarity between them (i.e., probabilistic), in which the meanings of words bring them closer together in that space in terms of statistical distances. It is folded in a dimension space. Its pet name: is word2vec or a language model in a more general way.

Language models, introduced in 1998 by Ponte and Croft [36] are a probabilistic concept. What is the probability that a query (sequence of terms (or grams)) will appear in a given document? The difficulty is establishing this model for small documents. In order to solve the lack of information during the construction of the model, different smoothing methods (Hiemstra, 98) (Song et al., 99) and different models (Berger et al., 1999) have been proposed. The results obtained by these models have shown equivalent or even better performances than the classical models. However, the term mismatch problem occurs frequently in IR. It can occur when the query is short and/or ambiguous but also in specialized domains where non-specialists make queries and documents are written by experts. Recently, the term mismatch problem has been addressed using neural learning to rank models and word plunging to avoid using only exact term matches for search. Another approach to the term mismatch problem is to use Knowledge Bases that can associate different terms with the same concept. In addition, the recent success of transformers in automatic natural language processing (NLP) [37][38][39][40], which have managed to achieve significant performance gains through sequence representations useful for this field of application. However, applying these encoders, for IR, in an architecture adopting deep learning does not lead to the same performance gain. The consulted literature presents the following explanations:

- The semantic matching used for NLP differs from the relevance matching adopted for IR. These differences affect the design of deep model architectures, and it can be difficult to find a "one-size-fits-all" solution to such different matching problems. [25][28]
- The query is usually short and based on keywords. The document can vary considerably in length, from tens of words to thousands or even tens of thousands of words.
- Deep learning techniques have been widely criticized as a "black box" that produces good results but no insights and explanations of problems.

We are nearing the end of this study, but we are still far from the end of the road in this line of research; there are still many open questions, unexplored directions, and much more work to do.

9. CONCLUSION

In this paper, we have presented some approaches to TIR. Five textual IR approaches have been discussed, those based on keywords, those based on graphs, those based on classification techniques, and those based on deep learning techniques. The traditional approaches suffer from polysemy and synonymy, while the language model-based approaches are more efficient because they allow for term linking (word-word). In addition, approaches based on ML techniques, in general, can encompass different approaches (statistical, graph, language...) on the one hand and benefit from several advantages, namely: feature extraction, system abstraction, response time, scalability, adaptability, etc. Hence the importance of investing in the advantages of the proposed approaches to have a more optimal model. The combination of several approaches is excellent potential to improve the performance of the IRS.

In our future work, we will propose a hybrid approach to IR based on the combination of several approaches of IR and deep learning techniques.

10. REFERENCES

[1] N. J. Belkin, “Helping people find what they don’t know: recommendation systems help users find the corret words for a successful search”, Communications of the ACM, Vol. 43, No. 8, 2000, pp. 58-61.

[2] B. Kitchenham, S. Charters, “Методи за автоматично управление на подемни устройства при Jack-up системите”, 2007.

[3] G. Salton, “Search and Retrieval Experiments in Real-Time Information Retrieval”, Technical Report, Cornell University, 1968, p. 16.

[4] B. A. Kitchenham, “Systematic review in software engineering: where we are and where we should be going”, Proceedings of the 2nd international workshop on Evidential assessment of software technologies, 2012, pp. 1-2.

[5] H. Benhar, A. Idri, J. L Fernández-Alemán, "Data preprocessing for heart disease classification: A systematic literature review", Computer Methods and Programs in Biomedicine, Vol. 195, 2020.

[6] S. Qaiser and R. Ali, “Text Mining: Use of TF-IDF to Examine the Relevance of Words to Documents”, International Journal of Computer Applications, Vol. 181, No. 1, 2018, pp. 25-29.

[7] G. Salton, J. Allan, C. Buckley, "Approaches to passage retrieval in full text information systems", Proceedings of the 16th annual international ACM SIGIR conference on Research and development in information retrieval, pp. 49-58, 1993.

[8] J. Ramos, "Using TF-IDF to Determine Word Relevance in Document Queries", Proceedings of the First International Conference on Machine Learning, 2003, pp. 29-48.

[9] O. A. S. Ibrahim and D. Landa-Silva, "Term frequency with average term occurrences for textual information retrieval", Soft Computing, Vol. 20, No. 8, 2016, pp. 3045-3061.

[10] R. B.-Y. WB Frakes, "Information retrieval: data structures and algorithms", Pearson College Div, 1992.

[11] M. Gamon, "Graph-based text representation for novelty detection", Proceedings of TextGraphs: the First Workshop on Graph Based Methods for Natural Language Processing, June 2020, pp. 17-24.

[12] R. Blanco, C. Lioma, "Graph-based term weighting for information retrieval", Information Retrieval Journal, Vol. 15, No. 1, 2012, pp. 54-92.

[13] M. R. Islam, M. R. Islam, "An improved keyword extraction method using graph based random walk model", Proceedings of the 11th International Conference on Computer and Information Technology, 2008, pp. 225-229.

[14] O. Sornil, "A Graph-Based Information Retrieval Model", Proceedings of the Conférence en Recherche d'Information et Applications, 2008.

[15] H. Patil, R. S. Thakur, "Document Clustering", Information Retrieval and Management: Concepts, Methodologies, Tools, and Applications, 2016, pp. 264-281.

[16] C. Bouveyron, S. Girard, "Classification supervisée et non supervisée des données de grande dimension", La revue MODULAD, Vol. 40, 2009, pp. 81-102.

[17] W. Lam, M. Ruiz, P. Srinivasan, "Automatic text categorization and its application to text retrieval", IEEE Transactions on Knowledge and Data Engineering, Vol. 11, No. 6, 1999, pp. 865-879.

[18] B. Trstenjak, S. Mikac, D. Donko, "KNN with TF-IDF based framework for text categorization", Procedia Engineering, Vol. 69, 2014, pp. 1356-1364.

[19] A. Desai, C. Shrihari, "A Review on Knowledge Discovery using Text Classification Techniques in Text Mining", International Journal of Computer Applications, Vol. 111, No. 6, pp. 12-15, 2015.

[20] J. Han, M. Kamber, J. Pei, "Data Mining Concepts and Techniques", Journal of Chemical Information and Modeling, Vol. 53, No. 9, 2012, pp. 1689-1699.

[21] S. S. Deeksha, "Finding similarity in articles using various clustering techniques", Proceedings of the 6th International Conference on Reliability in Information Communication Technology, Vol. 2018, June 2018, pp. 343-347.

[22] J. Lin, R. Nogueira, A. Yates, D. R. Cheriton, C. Science, "Pretrained Transformers for Text Ranking BERT and Beyond", Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval, 2021, pp. 1-204.

[23] F. Chollet, "Deep Learning with Python", Manning, 2017.

[24] S. E. Robertson, S. Walker, "Some Simple Effective Approximations to the 2-Poisson Model for Probabilistic Weighted Retrieval", Proceedings of the 17th annual international ACM SIGIR conference on Research and development in information retrieval, 1994.

[25] J. Guo, Y. Fan, Q. Ai, W. B. Croft, "A deep relevance matching model for Ad-hoc retrieval", Proceedings of the 25th ACM International on Conference on Information and Knowledge Management, 24-28-October 2016, pp. 55-64.

[26] P. Huang et al. "Learning Deep Structured Semantic Models for Web Search using Clickthrough Data", Proceedings of the 22nd ACM international conference on Information & Knowledge Management, 2013, pp. 2333-2338.

[27] Y. Shen, X. He, J. Gao, L. Deng, G. Mesnil, "A latent semantic model with convolutional-pooling structure for information retrieval", Proceedings of the 23rd ACM International Conference on Conference on Information and Knowledge Management, November 2014, pp. 101-110.

[28] J. Guo, Y. Fan, L. Pang, L. Yang, Q. Ai, "A Deep Look into neural ranking models for information retrieval", Information Processing & Management, No. June, 2019, p. 102067.

[29] H. Palangi et al. "Deep Sentence embedding using long short-term memory networks: Analysis and application to information retrieval", IEEE/ACM Transactions on Audio, Speech, and Language Processing, Vol. 24, No. 4, pp. 694-707, 2016.

[30] K. Hui, A. Yates, K. Berberich, G. de Melo, "PACRR: A position-aware neural IR model for relevance matching", Proceedings of the Conference on Empirical Methods in Natural Language Processing, July 2017, pp. 1049-1058.

[31] A. Yates, L. Sun, J. Xu, "NPRF: A Neural Pseudo Relevance Feedback Framework for Ad-hoc Information Retrieval", Proceedings of the Conference on Empirical Methods in Natural Language Processing, 2018, pp. 4482-4491.

[32] A. Berger, R. Caruana, D. Cohn, D. Freitag, V. Mittal, "Bridging the lexical chasm: statistical approaches to answer-finding", Proceedings of the 23rd annual international ACM SIGIR conference on Research and development in information retrieval, 2000, pp. 192-199.

[33] B. Trstenjak, S. Mikac, D. Donko, "KNN with TF-IDF based framework for text categorization", Procedia Engineering, Vol. 69, 2014, pp. 1356-1364.

[34] D. Beeferman, A. Berger, J. Lafferty, "Statistical Models for Text Segmentation", Machine Learning, Vol. 210, 1999, pp. 177-210.

[35] Z. Zhang, Z. Guo, C. Faloutsos, E. P. Xing, J. Y. Pan, "On the scalability and adaptability for multimodal retrieval and annotation", Proceedings of the 14th International Conference of Image Analysis and Processing - Workshops, Modena, Italy, 10-13 September 2007, pp. 39-44.

[36] J. M. Ponte, W. B. Croft, "A Language Modeling Approach to Information Retrieval", Proceedings of the 21st annual international ACM SIGIR conference on Research and development in information retrieval, 1998, pp. 275-281.

[37] M. C. Kenton, L. Kristina, J. Devlin, "BERT: Pretraining of Deep Bidirectional Transformers for Language Understanding", Proceedings of the Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers), 1953, pp. 4171-4186.

[38] G. Lample, A. Conneau, "Cross-lingual Language Model Pretraining", Proceedings of the Advances in Neural Information Processing Systems 32, 2018.

[39] A. Vaswani, "Attention Is All You Need", Proceedings of the 31st Conference on Neural Information Processing Systems, Long Beach, CA, USA, 2017.

[40] A. Radford, J. Wu, R. Child, D. Luan, D. Amodei, I. Sutskever, "Language Models are Unsupervised Multitask Learners", OpenAI, Whitepaper, 2018.

Low cost microcontroller implementation of Takagi–Sugeno Fuzzy MPPT controller for photovoltaic systems

Case Study

M'hand Oubella
Ibn Zohr University,
ESTA, Electrical Engineering Department,
Laboratory of Engineering Sciences and Energy Management (LASIME),
ESTA, BP: S33 – 80150, Agadir, Morocco
m.oubella@uiz.ac.ma

Mohamed Ajaamoum
Ibn Zohr University,
ESTA, Electrical Engineering Department,
Laboratory of Engineering Sciences and Energy Management (LASIME),
ESTA, BP: S33 – 80150, Agadir, Morocco
m.ajaamoum@uiz.ac.ma

Brahim Bouachrine
Ibn Zohr University,
ESTA, Electrical Engineering Department,
Laboratory of Engineering Sciences and Energy Management (LASIME),
ESTG, BP: S1317 – 81000, Guelmim, Morocco
b.bouachrine@uiz.ac.ma

Abstract *– Maximum power point trackers have a significant role in optimizing the energy conversion efficiency in a photovoltaic system. The numeric achievements of MPPT algorithm can be implemented and tested by using several embedded boards as Digital Signal Processor, Field-Programmable Gate Array, Arduino, and dspace. Alternatively, for the low cost, availability and simplicity, the PIC microcontrollers can be used compared with the other hardware devices. Therefore, this paper presents the implementation of a Takagi–Sugeno fuzzy controller on a low cost PIC microcontroller, for tracking the maximum power point of a PV module. The PV system consists of a PV emulator, DC-DC converter, and resistive load. The different steps in design, simulation and realization of the T-S Fuzzy logic controller are discussed. The proposed controller system was evaluated by comparing its performance metrics, in terms of efficiency and the Integral Square Error, with existing technique in the literature. The results corresponding to the experimental validation show that the proposed MPPT controller is able to ensure a perfect tracking of the maximum power point with variation of irradiance and is performing better than reported in a previous work.*

Keywords*: Photovoltaic, MPPT, Fuzzy Logic, Takagi–Sugeno, PIC16F877*

1. INTRODUCTION

For the past decades, the maximum power point tracking (MPPT) technique, in photovoltaic (PV) systems, has received a lot of attention in order to optimize the performance in these systems. The maximum power point (MPP) is not static; rather, it varies according to changes in atmospheric conditions. Therefore, MPPT is adopted to track and maintain the optimal operating condition of PV system, regardless of changing environment conditions [1]. MPPT techniques are most commonly applied in the DC-DC converters, which are used in photovoltaic systems as the MPPT circuit. The performance parameters of the PV system depend on both these converters and the MPPT algorithms. Many MPPT algorithms have been developed and implemented by different researchers [2], [3] and have offered various explanations for the problems associated with the MPPT controller. Recently, several comparative studies, of these algorithms, have been carried out in [4]–[6], where the MPPT algorithms are

categorized based on three main features: complexity, performance, and cost of implementation.

In literature, the conventional MPPTs are mostly used due to their ease of implementation. Among these MPPTs are Perturb and Observe [7], Hill Climbing [8], and Incremental Conductance [9]. Nevertheless, the main drawback of these algorithms is the power oscillation in the vicinity of the MPP, which leads to power loss and increases convergence time of the algorithm [4].

Several modern methods based on artificial intelligence have been proposed to remedy these issues, such as Artificial Neural Networks [10], and fuzzy logic controller (FLC) based *Mamdani or Takagi-Sugeno* model [11], [12]. These methods are accurate and efficient compared to the conventional techniques. Moreover, they respond quickly to changing environmental conditions and does not requiring any information of the PV model [4]. In order to design an accurate system, FLC uses knowledge and experience to construct the fuzzy rules base.

The achievement of MPPT algorithm can be implemented and tested by using microcontrollers [13], or embedded boards such as Digital Signal Processor [14], Field-Programmable Gate Array FPGA [15], Arduino [16], dSPACE [17]. For the implementation of MPPT algorithms, a proposal of the most appropriate embedded board was presented and discussed in [18]; the study was only focused on the low-cost electronic board. However, authors did not mention the cost-effective PIC microcontroller. Therefore, in this paper, we aim to contribute to the design and implementation of T-S fuzzy logic MPPT controller for a PV system on low cost microcontroller of type PIC 16F877. The hardware implementation of the MPPT algorithm based T-S Fuzzy Logic is carried out and the experimental prototype is tested for the PV system. The performance of the proposed controller is compared with the previously reported technique in the literature.

The remaining sections are organized as follows: in section 2, the investigation of PV system is presented. The section 3 is reserved to present the MPPT controller. Section 4 is devoted to hardware design and software implementation. Then, section 5 presents the experimental results and discussion on the effectiveness of the proposed MPPT controller. Finally, the conclusions are drawn in section 6.

2. PHOTOVOLTAIC SYSTEM AND MPPT CONTROLLER

Fig. 1 shows the system configuration used in this paper. This system is made up of a PV generator (PVG), current and voltage sensors, a DC-DC buck-boost converter, and a resistive load. Once voltage and current of the PV module are measured, they are sent to the control system, in which the algorithms are embedded. The duty cycle value of the buck-boost converter is adjusted by the MPPT based on T-S FLC and amplified by a driver circuit to get maximum power.

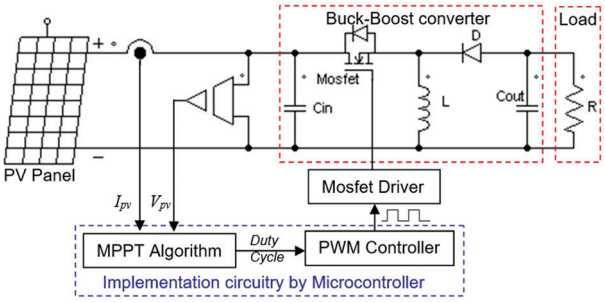


Fig. 1. Circuit diagram of PV system with MPPT based on T-S Fuzzy controller

2.1 PV PANEL MODEL

The PV panel is composed of several PV cells connected in series or parallel or both. The equivalent circuit of the single-diode model for PV cells is shown in Fig. 2. By neglecting the voltage drop in the series resistance of the generator and the current in the shunt resistor, the commonly used expression of the output current I_{PV} and output voltage V_{PV} of a PV panel with N_S cells in series and N_p strings in parallel [19] is:

$$I_{pv} = N_p\left(I_{ph} - I_0\left(exp\left(\frac{V_{pv}}{N_s V_T}\right) - 1\right)\right) \quad (1)$$

Where: I_{ph}, I_0 and V_T denote respectively the photocurrent, the reverse saturation current of the diode and the thermal voltage.

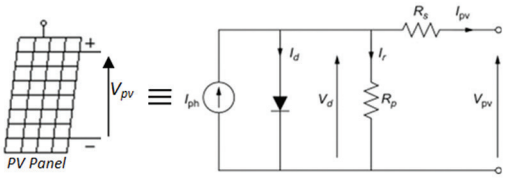


Fig. 2. Equivalent circuit for PV cell

In this work, a didactic bench emulating a solar module "CO3208-1A" (Fig. 3) is employed to simulate a PVG [20]. This bench consists of three independent blocks emulating solar panels; each with adjustable irradiance and integrated voltage and current displays.

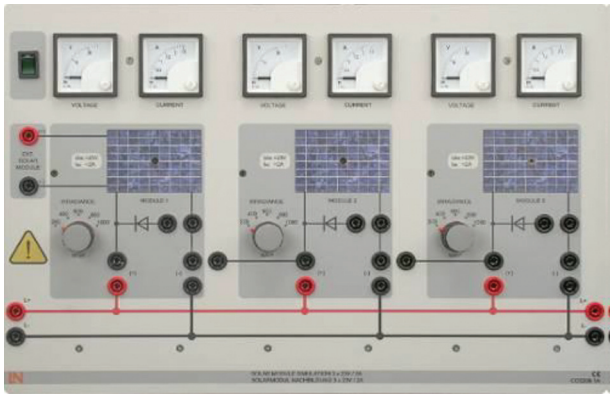


Fig. 3. "CO3208-1A" didactic bench emulating a solar module [20]

The key specifications of each block of this module are shown in table 1 under standard test conditions (STC), i.e., 25 °C and 1000W/m².

Table 1. The specifications of the solar module [20]

Model Parameters	Specification CO3208-1A
Open circuit voltage (V_{oc})	23V
Short circuit current (I_{sc})	2A
Maximum power (P_{mpp})	38W
Voltage at maximum power (V_{mpp})	21V
Current at maximum power (I_{mpp})	2A

Electrical characteristics of PV module Current-Voltage (I-V) and Power-Voltage (P-V) are illustrated by Fig. 4. It is very clear that these characteristics have non linear behavior with a single point where the PV power is maximum. This point varies with irradiation. Therefore, it is crucial to control DC-DC converter switch periodically via MPPT controller to have optimal energy of PV panel.

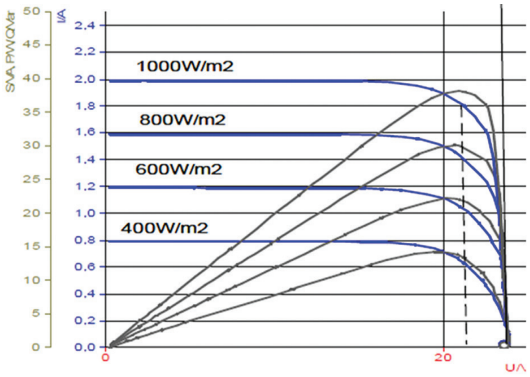


Fig. 4. The characteristics I-V and P-V of the solar module "CO3208-1A" [20]

2. 2 BUCK-BOOST CONVERTER

Buck-Boost converter, shown in Fig. 1, is chosen in the present work for matching the characteristics of the load with those of the PV panel. This converter is controlled using a Pulse-Width Method (PWM) and the duty cycle α is calculated for tracking the MPP of the PV panel.

The Buck-Boost converter can operate as Buck (for voltage step down) or Boost (for voltage step up) converter. However, the output voltage has the opposite sign to the input voltage. There are two operating modes for the buck-boost converter, namely, the continuous conduction mode (CCM) and the discontinuous conduction mode (DCM). Equations (2) to (4) are used to modeling buck-Boost converter in Continuous Conduction Mode.

$$\alpha = \frac{V_{out}}{V_{in} + V_{out}} \tag{2}$$

$$L = \frac{V_{in}.\alpha}{f_s.\Delta I_L} \tag{3}$$

$$C = \frac{V_{out}.\alpha}{f_s.\Delta V_{out}.R} \tag{4}$$

Where V_{in} is the input voltage, V_{out} is the output voltage and α is the duty ratio determined by the applied MPPT algorithm to directly control the converter switching. Δi_L is the current ripple, L is the inductance, ΔV_{out} is the output voltage ripple, C is the output capacitor, R is the load resistance, and f_s is the switching frequency.

2. 3 MAXIMUM POWER POINT TRACKING

According to the P-V characteristic of the PV panel shown in Fig. 4, the maximum power is produced for a particular voltage condition that depends on solar irradiation and temperature. Consequently, the MPP changes by moving on the PV curve depending on the weather conditions. As a result, an MPPT technique is needed so that the produced energy is always maintained at its maximum.

The PV panel produces a power expressed as follows: $P_{pv} = I_{pv}.V_{pv}$, hence:

$$P_{pv} = N_p\left(I_{ph} - I_0\left(exp\left(\frac{V_{pv}}{N_s V_T}\right) - 1\right)\right)\cdot V_{pv} \tag{5}$$

The condition which must be satisfied for the maximum power point PPM is:

$$\frac{dP_{pv}}{dV_{pv}} = 0 \tag{6}$$

$$\frac{dP_{pv}}{dV_{pv}} = I_{pv} - V_{pv}\frac{N_P.I_0}{N_s.V_T}exp\left(\frac{V_{pv}}{N_s V_T}\right) \tag{7}$$

The nonlinearity of equation (7) makes MPPT search difficult, so it is not easy to come by the MPPT using traditional methods. Thus, we will use the T-S fuzzy method to find the maximum power point.

3. MPPT T-S FUZZY CONTROLLER

3.1 CONCEPT OF MPPT FUZZY LOGIC CONTROLLER

The power–voltage characteristic illustrated in Fig. 5 shows that the variation of the maximum power (P_{MPP}) of PVG, which corresponds to an optimal PV output voltage (V_{MPP}), can be changed as a function of the solar irradiation and cell temperature. We introduce the concept of MPPT FLC as follows: The output power of PVG is examined by the FLC in each sample time (t) and then defines the change in power with respect to voltage (dp_{pv}/dv_{pv}). If this value is bigger than zero, the controller modifies the duty cycle $\Delta\alpha$ of PWM signal controlling the DC-DC converter to increase the voltage until the power is maximum ($dp_{pv}/dv_{pv} = 0$). If the value (dp_{pv}/dv_{pv}) is lower than zero the controller changes the duty cycle Δα to reduce the voltage until the power is maximized and so on. Fig. 5 illustrates this concept.

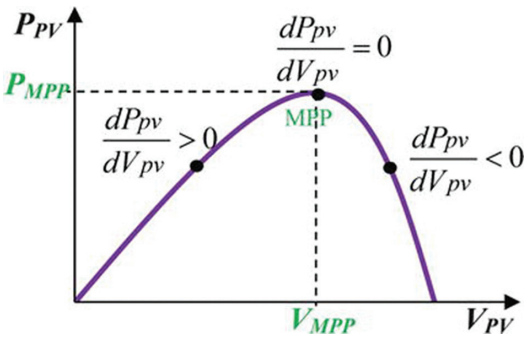


Fig. 5. The concept of MPPT fuzzy logic controller

3.2 THE PROPOSED STRUCTURE OF THE T-S FUZZY CONTROLLER

Two types of Fuzzy Inference Systems (FIS), *Mamdani and Takagi-Sugeno* are used in control applications. The *Takagi and Sugeno* (T-S) fuzzy model can be reformulated as a fuzzy dynamic model to represent nonlinear dynamic systems [21]. This fuzzy modelling method provides an alternative approach to describing complex nonlinear systems, compared with *Mamdani* fuzzy model; T-S fuzzy model can reduce the number of rules in modeling higher-order nonlinear systems [22]. The two models differ in compilation of output. The consequence parts of the *Mamdani* fuzzy model are fuzzy sets while those of the T-S fuzzy model are linear functions of input variables. In this paper, maximum power point tracking is implemented using a T-S fuzzy controller.

The principal operating of T-S Fuzzy inference systems is shown in Fig. 6.

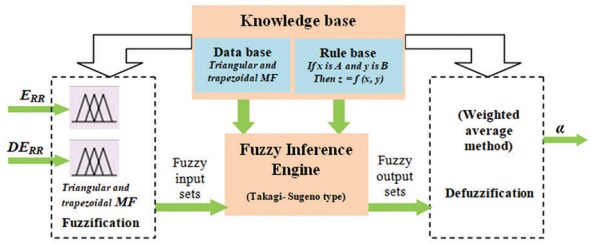


Fig. 6. T-S fuzzy model schema diagram principle

The fuzzy inference system contains five functional components:

- A fuzzification interface is used to convert crisp inputs (E_{RR} and DE_{RR}) to linguistic variables as fuzzy inputs applying membership functions (*MF*).
- A database is used to define membership functions.
- A rule base contains several fuzzy if-then rules.

The T-S style fuzzy rule is: *IF x is A AND y is B THEN z is f (x, y)* where *x, y* and z are linguistic variables, *A* and *B* are fuzzy sets on universe of discourses *X* and *Y*, and *f (x, y)* is a mathematical function.

- The inference engine of the FLC is a decision-making unit.
- The defuzzification interface is used to convert fuzzy outputs into crisp outputs.

The structure of the proposed T-S fuzzy controller is characterized by:

- The error E_{RR} (t) and change in error DE_{RR} (t) inputs, at sampled times t, are calculated by:

$$E_{RR}(t) = \frac{\Delta P_{pv}}{\Delta V_{pv}} = \frac{P_{pv}(t) - P_{pv}(t-1)}{V_{pv}(t) - V_{pv}(t-1)} \tag{8}$$

$$DE_{RR}(t) = E_{RR}(t) - E_{RR}(t-1) \tag{9}$$

- $P_{pv}(t)$ and $V_{pv}(t)$ are respectively the instantaneous power and voltage delivered by the PVG.

The membership functions for input error ERR are defined as:

$$\mu_i(NB) = Max\left(min\left(\left[\frac{-a_1 - E_{RR}}{-a_1 + a_2}\right], 1\right), 0\right) \tag{10}$$

$$\mu_i(NM) = Max\left(min\left(\left[\frac{E_{RR} + a_2}{-a_1 + a_2}\right], \frac{-E_{RR}}{a_1}\right), 0\right) \tag{11}$$

$$\mu_i(ZE) = Max\left(min\left(\left[\frac{E_{RR} + a_1}{a_1}\right], \left[\frac{a_1 - E_{RR}}{a_1}\right]\right), 0\right) \tag{12}$$

$$\mu_i(PM) = Max\left(min\left(\frac{E_{RR}}{a_1}, \left[\frac{a_2 - E_{RR}}{a_2 - a_1}\right]\right), 0\right) \tag{13}$$

$$\mu_i(PB) = Max\left(min\left(\left[\frac{E_{RR} - a_1}{a_2 - a_1}\right], 1\right), 0\right) \tag{14}$$

The change in error DERR membership functions is defined by the same Equations above (10 - 14) by changing a_1, a_2 and E_{RR} to b_1, b_2 and DE_{RR} respectively.

- The single output variable is the change in duty cycle $\Delta\alpha$ of PWM signal controlling the DC-DC converter.

$\Delta\alpha$ is defined by the control law, so that

$$\Delta\alpha(t) = (E_{RR}, DE_{RR})$$

The duty cycle α is defined as:

$$\alpha(t) = \alpha(t-1) + \Delta\alpha(t) \tag{15}$$

In the Takagi-Sugeno modeling, the final output is equal to the weighted average of the output of each rule. By using the Equation (16), we evaluate the net value $\Delta\alpha$ of the output variable, in which the summations are based on the 25 rules.

$$\Delta\alpha = \frac{\sum_{i=1}^{25} Z_i w_i}{\sum_{i=1}^{25} w_i} \tag{16}$$

With Z_i is the output level of each rule R_i and w_i is the degree of activation of each rule calculated by the following expression:

$$w_i = \mu_{E_{RRi}} \bullet \mu_{DE_{RRi}} \tag{17}$$

- The universe of discourse, which is normalized to the interval [-1, 1], is spitted into five fuzzy sets for input and output variables.

The triangular and trapezoidal membership functions are used in the proposed T-S fuzzy controller for inputs and singleton for output as well. Fig. 7 illustrates the appearance of these functions.

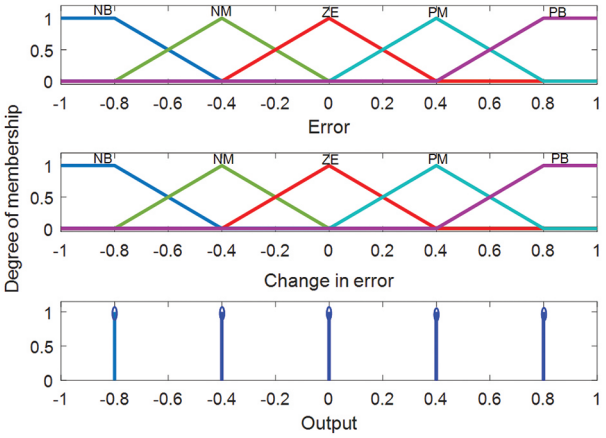


Fig. 7. Membership functions plots of the input variables (E_{RR}, DE_{RR}) and the output variable α

- Base of 25 rules.

We named the five fuzzy sets as follows: NB for negative big, NM for negative medium, ZE for about zero, PM for positive medium, and PB for positive big. Table 2 summarizes the basis of rules, which includes the 25 rules.

The flowchart in Fig. 8 graphically describes the steps of the algorithm required to generate a control signal on an output of the microcontroller. The output signal is connected to the MOSFET gate driver of the Buck-Boost converter.

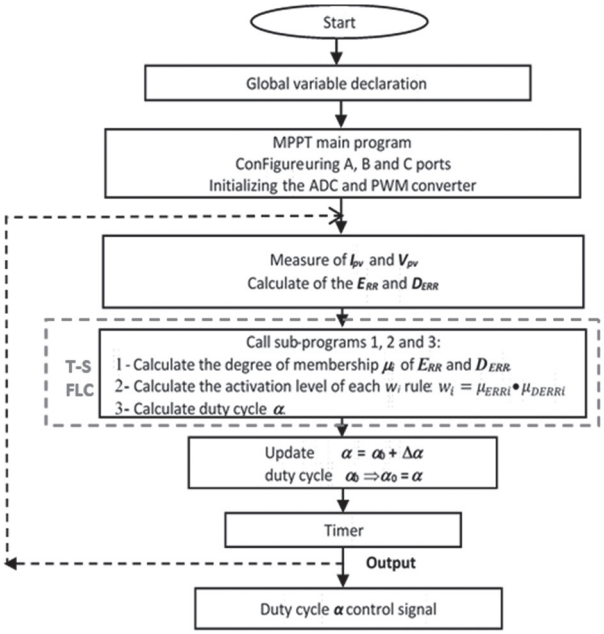


Fig. 8. Flowchart of MPPT algorithm based T-S Fuzzy logic

Table 2. Base of fuzzy rules

ERR / DERR	NB	NM	ZE	PM	PB
NB	PB	PB	PB	PM	ZE
NM	PB	PB	PM	ZE	NM
ZE	PB	PS	ZE	NM	NB
PM	PM	ZE	NM	NB	NB
PB	ZE	NM	NB	NB	NB

4. HARDWARE DESIGN AND SOFTWARE IMPLEMENTATION

This section offers details of the hardware components and the operation used in the experimental prototyping platform along with the software implementation. For the convenience of demonstration, we divided the section into two parts: software implementation and hardware description.

4.1. SOFTWARE IMPLEMENTATION OF MPPT ALGORITHM BASED T-S FUZZY LOGIC CONTROLLER

We translated the flowchart (Fig. 8) into a program implemented in the PIC 16F877 microcontroller through C-code. After several tests, we came up with a version of the program, which consists of three subprograms that describe the T-S Fuzzy Controller as illustrated in Fig. 8. These subprograms are detailed below.

First Sub-Program: The first sub-program, aims to calculate the degree of membership μ_i of the E_{RR} and DE_{RR}, for the five fuzzy classes used. The membership functions are calculated through the block diagram realized under Simulink (Fig. 9), which makes it easier to understand the program written in C language.

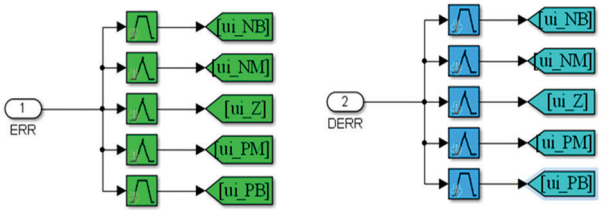


Fig. 9. Simulink model for fuzzification process

For the calculations of Equation (10) to (14), we are required to create *MAX* and *MIN* functions used to determine the degrees of membership for E_{RR} and DE_{RR} inputs, because the functions available for the PIC 16F877 microcontroller use integers (*Int*) and we use real numbers (*float*).

Second Sub-Program: With reference to Fig. 7 and based on the aforementioned universe of discourse, the purpose of this subprogram is to calculate the activation level w_i of each rule, using Simulink block diagram (Fig. 10).

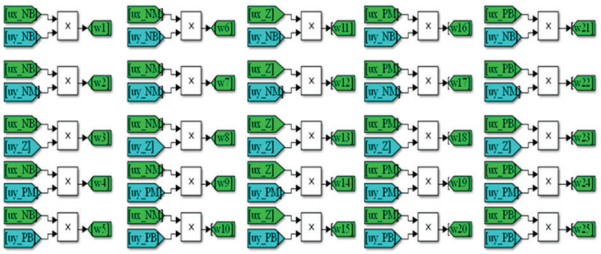


Fig. 10. Simulink model of the fuzzy rules

Third Sub-Program: After the rules application, a defuzzifier with the weighted average method is applied to adjust and generate the output analog signal. Thus, the call of the last sub-program allows calculating the new value of the duty cycle α that controls the switch of the Buck-Boost converter. Fig. 11 shows the blocks-set which achieve the expression of equations (15) and (16).

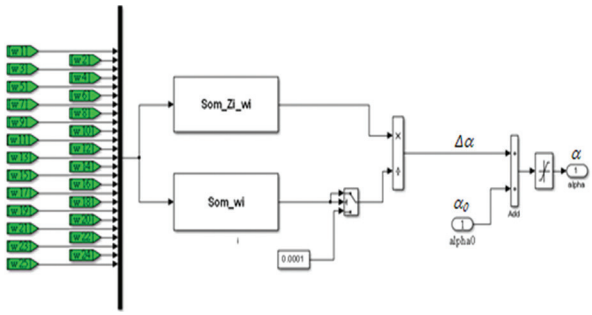


Fig. 11. Simulink model for defuzzification process

After obtaining satisfactory simulation results in MATLAB/Simulink, the Proteus 7 professional software was used for schematics drawing, circuits simulation and printed circuit board design. This software provides in its libraries electronic boards and microcontrollers such as PIC and Arduino. The PIC16F877 requires loading with a *HEX* file, through the use of the software tool MicroC PRO for PIC from Mikroelektronika. The schematic circuit of the embedded board developed for PIC16F877 is displayed in Fig. 12.

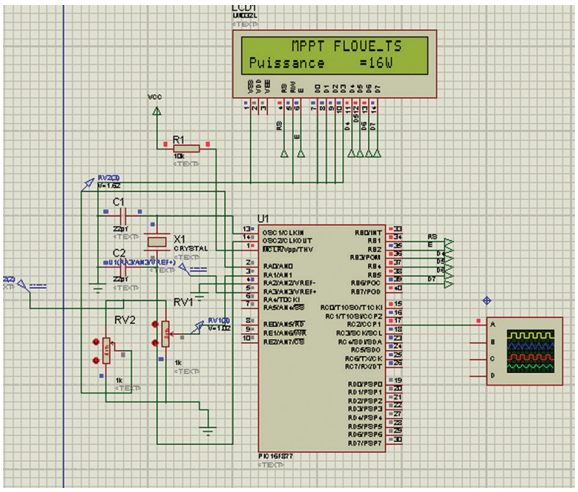


Fig. 12. Proteus schematic circuit developed using PIC16F877 Microcontroller

4. 2. HARDWARE DESCRIPTION

Fig. 13 below shows the block diagram of the hardware used for the experimentation corresponding to the photovoltaic system under study. It is composed of a PV emulator used as a power source, a Buck-Boost converter, I_{pv} and V_{pv} electrical sensors, and a resistive load. The PIC16F877 microcontroller is used to implement the proposed T-S Fuzzy algorithm.

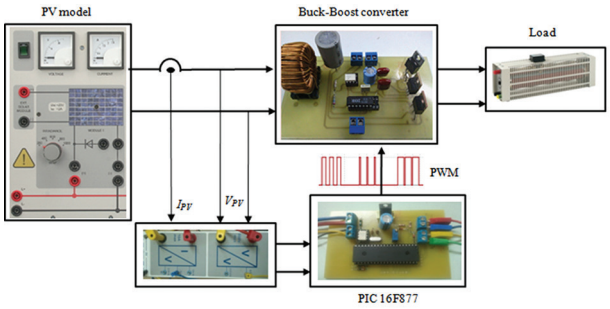


Fig. 13. Hardware implementation of the proposed T-S Fuzzy controller

The basic hardware components of the building blocks we used are included in the following sub-sections.

- *Buck-Boost converter*

The DC-DC converter is the main element of the PV system. It is an impedance matching circuit between the PVG and its load. A prototype of Buck-Boost converter shown in Fig. 14 is successfully developed to operate in the continuous current mode (CCM), according to the design steps developed in [23]. It is composed of a power MOSFET switch (IRF540), a Schottky diode (MBR20H), an inductor (2.75mH) and electrolytic filter capacitor (470µF) at the output side. The values of filter elements are calculated for the switching frequency of 15 KHz.

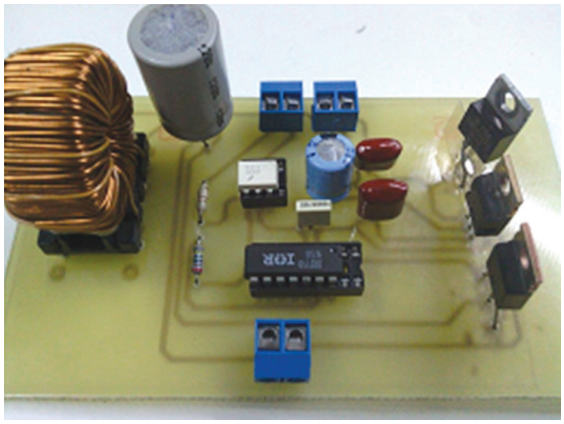


Fig. 14. Practical realization of the Buck-Boost converter.

The MOSFET transistor is controlled by a PWM signal generated from the MPPT control through an isolated Mosfet gate driver (IR2111) using simple optocoupler. The gate driver circuit is used to translate the voltage level of the PWM signal.

- *Current and voltage Sensors*

Fig. 15 illustrates a block used for measuring the current and voltage of the solar module. The block is equipped with a current sensor, based on a Hall-effect cell, and a voltage sensor; it consists of a model that incorporates sensors with the conditioning modules, which are arranged in an accessible form by printed connection terminals.

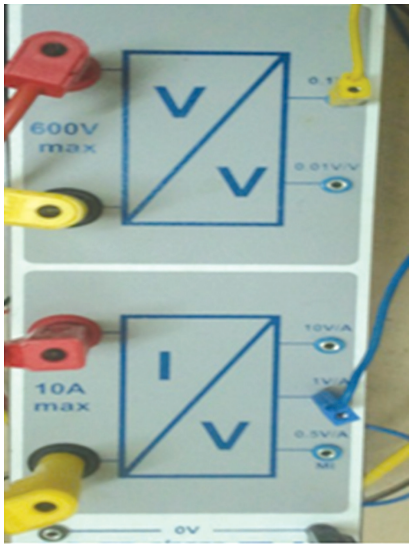


Fig. 15. The picture of the Current and voltage Sensors block.

The analogue outputs of sensors are given to on-chip Analog to Digital Converter (ADC) of the PIC16F877 microcontroller, which converts each analogue voltage to a decimal number thanks to the 10 bits resolution of the input ADC.

- *The PIC18F16778 Microcontroller Board*

Microcontroller is the common option of implementing the MPPT algorithm. The microcontroller selected, for our application, is PIC16F877 from Microchip Technology due to the availability of the device for the prototype. It is an 8-bit, 40-pins dual inline package (DIP). Table 3 describes key features of PIC16F877, which we used in this study.

Table 3. Key Features of the PIC16F877

Key Features	PIC16F877
Operating Frequency	20 MHz
Resets	POR, BOR
Flash Program Memory	8 K
Data Memory	368
EEPROM Data Memory	256
Interrupts	14
I/O Ports	Ports A,B,C,D,E
Timers	3
Capture/Compare/PWM modules	2
Serial Communication	USART
Parallel Communication	PSP
10 bit A/D module	8 input channels
Instruction set	35 instructions

Fig. 16 shows the embedded board developed for the PIC16F877, which represents the brain of the MPPT controller. It was designed and assembled using electronic circuits for programming. To generate the clock signal of the microcontroller, a 20 MHz crystal oscillator was used. The power supply of microcontroller board is considered basic and is, therefore, not discussed in this paper.

The microcontroller operation depends on the code that already programmed and uploaded into the board by using the software tool MicroC PRO for PIC, which provides a very efficient and advanced IDE (Integrated Development Environment).

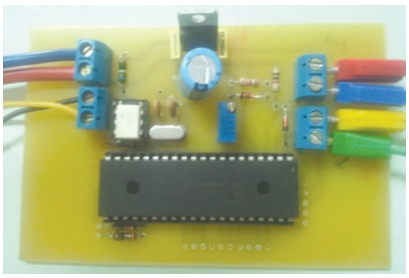


Fig. 16. The development board based on the PIC16F877 microcontroller

The output PWM signal is a square waveform with an amplitude value of 5 V at steady frequency of 15 kHz, and a variable duty cycle α based on the MPPT algorithm. The PWM signal feeds the MOSFET switch in the converter through the MOSFET driver.

5. EXPERIMENTAL RESULTS AND DISCUSSION

5.1. POWER–VOLTAGE CHARACTERISTIC OF THE MPPT BASED T-S FUZZY ALGORITHM

In this subsection, we consider the tracking performances under STC. To determine the system performance, we adopted three distinctive levels of irradiance G (600 W/m^2, 800 W/m^2 and 1000 W/m^2).

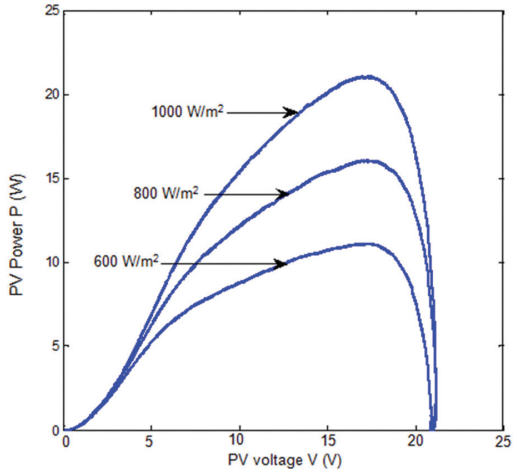


Fig. 17. P_{pv} (V_{pv}) characteristic of the Solar Module for different irradiation

Before launching the T-S MPPT controller, it is first necessary to carry out the plot of the static characteristic P_{pv} (V_{pv}), as a function of irradiation G, of the solar module used in this study. This characteristic is obtained using the block diagram of the hardware shown in Fig. 13 and the Matlab/Simulink environment. The Experimental P_{pv} (V_{pv}) characteristic (Fig.17) is provided to validate the performance of the proposed MPPT controller implemented on the PIC microcontroller.

Fig. 18 shows the experimental results of the MPPT tracking operation for different irradiations. We noticed that, due to the T-S Fuzzy MPPT algorithm implemented on PIC microcontroller, the MPP is reached with very little oscillations. These oscillations are occurred at the vicinity of MPP leading to power losses, which are almost neglected [4]. Hence, the proposed MPPT controller "T-S Fuzzy" is able to ensure a perfect tracking of the PPM of a solar module under variable irradiance.

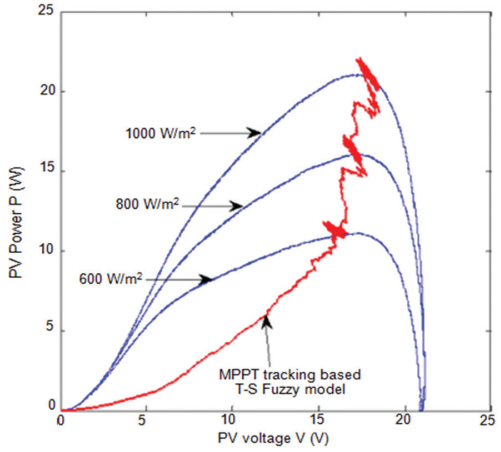


Fig. 18. MPPT tracking for different irradiation

The Tests of the MPPT controller "T-S Fuzzy", implemented on PIC microcontroller, are realized in static and dynamic conditions. Fig. 19 depicts the PV power for T-S Fuzzy MPPT technique for step-change in the irradiance from 600W/m² to 1000W/m².

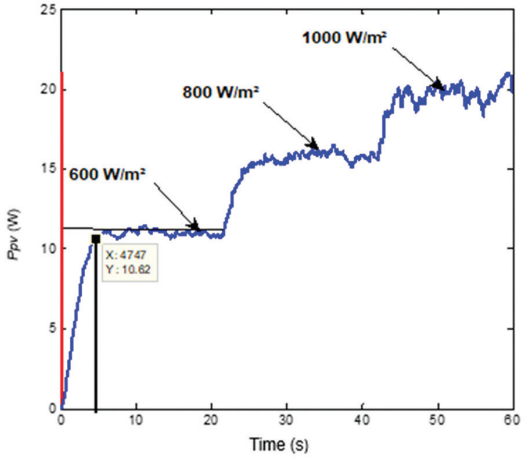


Fig. 19. Output power versus time under T=25 °C and step increasing of irradiance G

According Fig. 19, the respective maximum values of output power were 11.01, 16.13 and 20.35 W. These values are extremely close to the maximum values of the PV module as deduced from the experimental P_{pv} (V_{pv}) characteristics (Fig. 18).

Moreover, for the dynamics of the MPPT controller, it is noted from Fig. 19 that the output power reaches 96% of the MPP corresponding to G = 600 W/m² in 4.47 s. Based on these findings, we conclude that the proposed T-S Fuzzy MPPT controller achieves the correct tracking of the MPP, consequently proving its good performance in terms of finding the MPP under sudden changes in solar irradiation.

5.2. PERFORMANCES OF THE T-S FUZZY MPPT CONTROLLER IMPLEMENTED ON PIC MICROCONTROLLER

- *Comparison of Output Power of Solar Module with and without MPPT:*

A comparison is done between the output power with the MPPT circuit and without the MPPT circuit. The load here is taken to be the same for both circuits that is 30 Ω.

In Table 4, the efficiencies of the Photovoltaic System are compared with and without the MPPT. The Table displays the output power at load and the maximum power that can be delivered by the solar module under different irradiance.

Table 4. Comparison of output power of solar module without and with MPPT based T-S Fuzzy

Irradiance G (W/m²)	P_{pv} at load without MPPT (W)	P_{pv} at load with MPPT based T-S Fuzzy (W)	Efficiency (η)	Efficiency (%)
600	10.92	11.01	0.09	0.8
800	11.71	16.13	4.42	37.74
1000	12.87	20.35	7.48	58.11

Based on the results presented in Table 4 and for the developed MPPT T-S Fuzzy controller, the PV output power achieves the values 11.01, 16.13 and 20.35W, for the respective G values of 600, 800 and 1000 W/m². This presents almost 100% of the maximum power that the solar module can provide under the same experimental conditions. Comparing these results with those we found without of the presence of an MPPT controller, we conclude that the MPPT T-S Fuzzy controller contributed to a power gain of up to 58%, in the case of G =1000 W/m².

- *Performance criteria*

The experimental evaluation of the MPPT controller performances, implemented on PIC microcontroller, is presented by two performance criteria.

- The performance criterion η_{MPPT}, which characterizes the efficiency of the MPPT controller, defined by Equation (18).
- The Integral of the Squared Error (ISE) criterion, which characterizes the speed of response of MPPT controller. This criteria is defined by Equation (19):

$$\eta_{MPPT} = \frac{\int_0^t P_{min}(t)dt}{\int_0^t P_{max}(t)dt} \quad (18)$$

$$ISE = \int_0^t \left[e(t)\right]^2 dt \quad (19)$$

The criteria η_{MPPT} and ISE are calculated by means of the Simulink block diagram as shown in Fig. 20 (*a*) and (*b*), which are developed from the Equations (16) and (17) respectively.

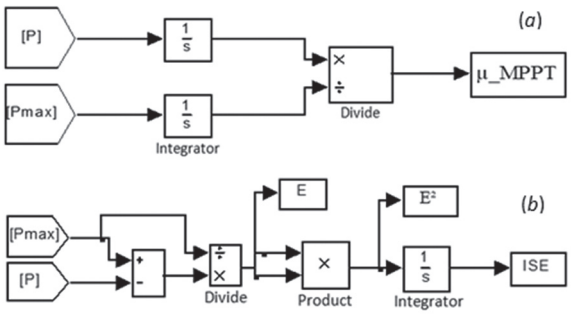


Fig. 20. Simulink model to determine: (*a*) Efficiency, (*b*) ISE.

Under the same experimental conditions that is for a step change of solar irradiance from 0 to 600W/m² and a period of 60s, we compare the performances of the two techniques used for the implementation of T-S Fuzzy MPPT controller; the rapid prototyping (T1) reported in [23] and the proposed technique (T2).

The results from comparing the criteria η_{MPPT} and ISE are presented in Figs. 21 and 22 respectively. The results of the proposed technique were obtained from the experimental measurements using the platform shown in Fig. 13, while those of the rapid prototyping are obtained from the experimental measurements using the PCI 1711 acquisition card [23].

As Fig. 21 shows, the curves associated with the two techniques are quite similar. However, we noticed that the T1 method tends to increase rapidly than T2. But, beyond 40 s, the trend is reversed, which enables reading the steady state faster with the proposed technique.

Under steady state condition, if the efficiency of an MPPT is greater than 97% then it is considered the best as reported in [24]. Hence, from the comparison of the η_{MPPT}, it follows that the proposed technique is quite effective, with its η_{MPPT} = 97.44% against η_{MPPT} = 96.32% for the rapid prototyping technique.

Under the same experimental conditions, Fig. 22 shows the "Integral of Squared Error" ISE curves determined for the two respective techniques T1 and T2. In this comparison, it is clear that the T2 technique is more efficient, with its ISE=1.28 lower than that of the T1 technique characterized by its ISE=1.44.

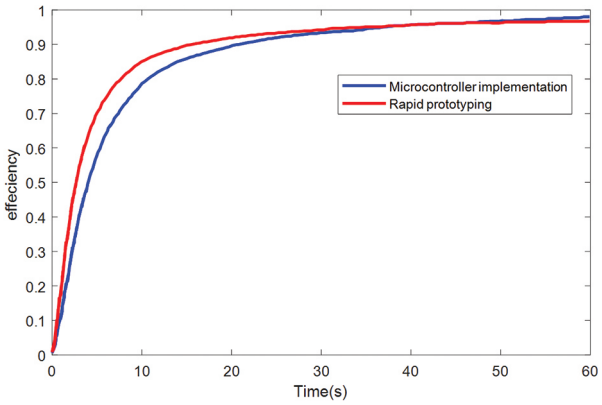


Fig. 21. η_{MPPT} criteria comparison

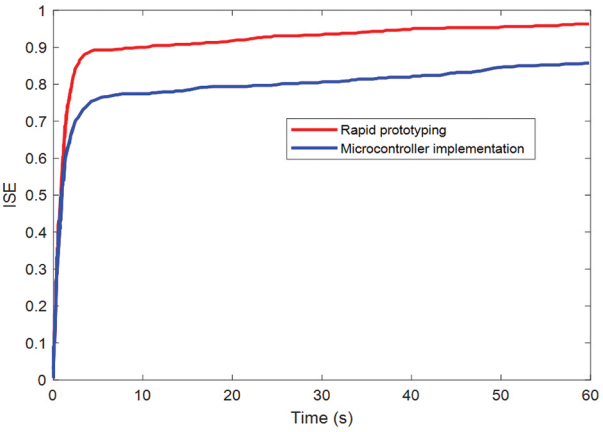


Fig. 22. ISE criteria comparison

Summary of the comparison between the two techniques

After presenting the results obtained by the realized T-S Fuzzy MPPT controller implemented on PIC microcontroller and compared with rapid prototyping reported in [23], we offer a summary of the results in Table 5. In this Table, we compared the two techniques based on three criteria: the efficiency η_{MPPT}, ISE Criteria and the response time.

Table 5. Comparison of parameters between rapid prototyping and implementation of MPPT "T-S Fuzzy" on PIC microcontroller

Technique	η_{MPPT} (%)	ISE	Response Time (s)
Rapid Prototyping [23]	96.32	1.44	2.39
Microcontroller implementation	97.44	1.28	4.74

Based on the results obtained in this work and the criteria previously cited, we can see that the proposed technique contributes to better results than the rapid prototyping in terms of efficiency and ISE criterion.

However, we notice that the response time of the proposed method is about two times greater than the rapid prototyping one. This difference is due to transient response of converter circuit and other circuitry. The low cost microcontroller as compared to the reported method compensates for this disadvantage, taking into account that the desired point to be reached in this system is obtaining the maximum yield from the PV panel.

6. CONCLUSION

This paper described the design of control strategy based on the *Takagi–Sugeno* fuzzy algorithm controller for the MPPT of a PV energy system. The T-S fuzzy algorithm has been successfully implemented in a low cost PIC 16F877 microcontroller. The proposed controller has been tested in a PV system consisting of a solar module, a Buck-Boost converter and a resistive load. Simulation and experimental results have been carried out to evaluate the effectiveness of the proposed controller system.

The results show that the MPPT controller "T-S Fuzzy" is able to ensure a perfect tracking of the PPM of a solar module under variable irradiance. The performance of the proposed controller is compared with the previously reported technique in the literature built around the PCI 1711 acquisition card in terms of efficiency and the Integral Square Error (ISE) index. The experimental results show the superiority of our T-S Fuzzy controller with η_{MPPT} = 97.44% against η_{MPPT} = 96.32% and ISE=1.28 lower than ISE=1.44 as compared with the existing techniques. Thus, the proposed controller based on low-cost microcontroller tends to perform better than controllers reported in previous work and can work properly for MPPT applications.

This work and the results presented herein, which can be used for PV energy harvesting as well, may be helpful for the design of the energy harvesting-based embedded systems. For further research, we intend to conduct a deeper experiment that corroborates our conclusions from this work by using more efficient MPPT algorithms to extract maximum power from PV system at real time.

7. REFERENCES

[1] M. A. Eltawil, Z. Zhao, "MPPT Techniques for Photovoltaic Applications", Renewable and Sustainable Energy Reviews, Vol. 25, 2013, pp. 793-813.

[2] Y. E. Abu Eldahab, N. H. Saad, A. Zekry, "Enhancing the maximum power point tracking techniques for photovoltaic systems", Renewable and Sustainable Energy Reviews, Vol. 40, 2014, pp. 505-514.

[3] M. S. Sivagamasundari, D. P. M. Mary, V. K. Velvizhi, "Maximum Power Point Tracking For Photovoltaic System by Perturb and Observe Method Using Buck Boost Converter", International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, Vol. 2, No. 6, 2013, p. 2433-2439.

[4] A. Gupta, Y. K. Chauhan, R. K. Pachauri, "A Comparative Investigation of Maximum Power Point Tracking Methods for Solar PV System", Solar Energy, Vol. 136, 2016, pp. 236-253.

[5] M. A. G. de Brito, L. Galotto, L. P. Sampaio, G. d. A. e Melo, C. A. Canesin, "Evaluation of the Main MPPT Techniques for Photovoltaic Applications", IEEE Transactions on Industrial Electronics, Vol. 60, No. 3, 2013, pp. 1156-1167.

[6] S. Motahhir, A. El Ghzizal, S. Sebti, A. Derouich, "MIL and SIL and PIL tests for MPPT algorithm", Cogent Engineering, Vol. 4, No. 1, 2017, p. 1378475.

[7] M. Abdel-Salam, M.T. El-Mohandes, M. Goda, "An Improved Perturb-and-Observe based MPPT Method for PV Systems under Varying Irradiation Levels", Solar Energy, Vol. 171, 2018, pp. 547-561.

[8] X. Xiao, X. Huang, Q. Kang, "A Hill-Climbing-Method-Based Maximum-Power-Point-Tracking Strategy for Direct-Drive Wave Energy Converters", IEEE Transactions on Industrial Electronics, Vol. 63, No. 1, 2016, pp. 257-267.

[9] R. I. Putri, S. Wibowo, M. Rifa'i, "Maximum Power Point Tracking for Photovoltaic Using Incremental Conductance Method", Proceedings of the 2nd International Conference on Sustainable Energy Engineering and Application, Bandung, Indonesia, 14-16 October 2014, pp. 22-30.

[10] M. Yaichi, M.K. Fellah, A. Mammeri, "A Neural Network Based MPPT Technique Controller for Photovoltaic Pumping System", International Journal of Power Electronics and Drive Systems, Vol. 4, No. 2, 2014, pp. 241-255.

[11] L. Suganthi, S. Iniyan, A. A. Samuel, "Applications of Fuzzy Logic in Renewable Energy Systems - A review", Renewable and Sustainable Energy Reviews, Vol. 48, 2015, pp. 585-607.

[12] M. Truntič, A. Hren, M. Milanovič, M. Rodič, "Adaptive Fuzzy-Logic State Controller for DC-DC Step-Down Converter", Electrical Engineering, Vol. 100, No. 2, 2018, pp. 983-995.

[13] B. R. S. Reddy, P. B. Narayana, P. Jambholkar, K. S. Reddy, "MPPT algorithm implementation for solar photovoltaic module using microcontroller", Proceedings of the Annual IEEE India Conference, Hyderabad, India, 16-18 December 2011, pp. 1-3.

[14] R. Akkaya, A. A. Kulaksız, Ö. Aydoğdu, "DSP Implementation of a PV System with GA-MLP-NN Based MPPT Controller Supplying BLDC Motor Drive", Energy Conversion and Management, Vol. 48, No. 1, 2007, pp. 210-218.

[15] A. Al-Gizi, B. Al-Rawe, M. Al-Saadi, A. Craciunescu, "Step by Step FPGA-Based Implementation of MPPT Fuzzy Controller for PV Systems", Proceedings of the 11th International Symposium on Advanced Topics in Electrical Engineering, Bucharest, Romania, 28-30 Mar. 2019, pp. 1-6.

[16] A. Al-gizi, M. Al-saadi, S. Al-chlaihawi, A. Craciunescu, M. A. Fadel, "Experimental Installation of Photovoltaic MPPT Controller Using Arduino Board", Proceedings of the International Conference on Applied and Theoretical Electricity, Craiova, Romania, 4-6 October 2018, pp. 1-6.

[17] O. F. Kececioglu, A. Gani, M. Sekkeli, "Design and Hardware Implementation Based on Hybrid Structure for MPPT of PV System Using an Interval Type-2 TSK Fuzzy Logic Controller", Energies, Vol. 13, No. 7, 2020, p. 1842.

[18] S. Motahhir, A. El Hammoumi, A. El ghzizal, "The Most Used MPPT Algorithms: Review and the Suitable Low-cost Embedded Board for Each Algorithm", Journal of Cleaner Production, Vol. 246, 2019, p. 118983.

[19] H. Abid, F. Tadeo, M. Souissi, "Maximum Power Point Tracking for Photovoltaic Panel based on T-S Fuzzy Systems", International Journal of Computer Applications, Vol. 44, No.22, 2012, pp. 50-58.

[20] L. Nülle, EPH 2.1 Investigating solar modules. https://www.lucas-nuelle.com/316/apg/13223/EPH-2.1-Investigating-solar-modules.htm (accessed: 2021)

[21] S. G. Cao, N. W. Rees, G. Feng, "Analysis and Design for a Class of Complex control Systems Part I: Fuzzy Modelling and Identification", Automatica, Vol. 33, No. 6, 1997, pp.1017-1028.

[22] M. Sugeno, "On Stability of Fuzzy Systems Expressed by Fuzzy Rules with Singleton Consequents", IEEE Transactions on Fuzzy Systems, Vol. 7, No. 2, 1999, pp. 201-224.

[23] M. Ajaamoum, M. Kourchi, B. Bouachrine, M. Oubella, A. Ihlal, L. Bouhouch, "Rapid Prototyping Test Bed of MPPT Photovoltaic Controller Based on Takagi-Sugeno Fuzzy Models", IREMOS, Vol. 12, No. 5, 2019, p. 303-312.

[24] R. Ahmad, A. F. Murtaza, H. A. Sher, "Power Tracking Techniques for Efficient Operation of Photovoltaic Array in Solar Applications - A review", Renewable and Sustainable Energy Reviews, Vol. 101, 2019, pp. 82-102.

INTERNATIONAL JOURNAL OF ELECTRICAL AND COMPUTER ENGINEERING SYSTEMS

Published by Faculty of Electrical Engineering, Computer Science and Information Technology Osijek,
Josip Juraj Strossmayer University of Osijek, Croatia.

About this Journal

The International Journal of Electrical and Computer Engineering Systems publishes original research in the form of full papers, case studies, reviews and surveys. It covers theory and application of electrical and computer engineering, synergy of computer systems and computational methods with electrical and electronic systems, as well as interdisciplinary research.

Topics of interest include, but are not limited to:

- Power systems
- Renewable electricity production
- Power electronics
- Electrical drives
- Industrial electronics
- Communication systems
- Advanced modulation techniques
- RFID devices and systems
- Signal and data processing
- Image processing
- Multimedia systems
- Microelectronics
- Instrumentation and measurement
- Control systems
- Robotics
- Modeling and simulation
- Modern computer architectures
- Computer networks
- Embedded systems
- High-performance computing
- Parallel and distributed computer systems
- Human-computer systems
- Intelligent systems
- Multi-agent and holonic systems
- Real-time systems
- Software engineering
- Internet and web applications and systems
- Applications of computer systems in engineering and related disciplines
- Mathematical models of engineering systems
- Engineering management
- Engineering education

Paper Submission

Authors are invited to submit original, unpublished research papers that are not being considered by another journal or any other publisher. Manuscripts must be submitted in doc, docx, rtf or pdf format, and limited to 30 one-column double-spaced pages. All figures and tables must be cited and placed in the body of the paper. Provide contact information of all authors and designate the corresponding author who should submit the manuscript to https://ijeces.ferit.hr. The corresponding author is responsible for ensuring that the article's publication has been approved by all coauthors and by the institutions of the authors if required. All enquiries concerning the publication of accepted papers should be sent to ijeces@ferit.hr.

The following information should be included in the submission:

- paper title;
- full name of each author;
- full institutional mailing addresses;
- e-mail addresses of each author;
- abstract (should be self-contained and not exceed 150 words). Introduction should have no subheadings;
- manuscript should contain one to five alphabetically ordered keywords;
- all abbreviations used in the manuscript should be explained by first appearance;
- all acknowledgments should be included at the end of the paper:
- authors are responsible for ensuring that the information in each reference is complete and accurate. All references must be numbered consecutively and citations of references in text should be identified using numbers in square brackets. All references should be cited within the text;
- each figure should be integrated in the text and cited in a consecutive order. Upon acceptance of the paper, each figure should be of high quality in one of the following formats: EPS, WMF, BMP and TIFF;
- corrected proofs must be returned to the publisher within 7 days of receipt.

Peer Review

All manuscripts are subject to peer review and must meet academic standards. Submissions will be first considered by an editor-in-chief and if not rejected right away, then they will be reviewed by anonymous reviewers. The submitting author will be asked to provide the names of 5 proposed reviewers including their e-mail addresses. The proposed reviewers should be in the research field of the manuscript. They should not be affiliated to the same institution of the manuscript author(s) and should not have had any collaboration with any of the authors during the last 3 years.

Author Benefits

The corresponding author will be provided with a .pdf file of the article or alternatively one hardcopy of the journal free of charge.

Units of Measurement

Units of measurement should be presented simply and concisely using System International (SI) units.

Bibliographic Information

Commenced in 2010.
ISSN: 1847-6996
e-ISSN: 1847-7003

Published: semiannually

Copyright

Subscription Information

The annual subscription rate is 50€ for individuals, 25€ for students and 150€ for libraries.

Postal Address

Faculty of Electrical Engineering,
Computer Science and Information Technology Osijek,
Josip Juraj Strossmayer University of Osijek, Croatia
Kneza Trpimira 2b
31000 Osijek, Croatia

IJECES Copyright Transfer Form

(Please, read this carefully)

This form is intended for all accepted material submitted to the IJECES journal and must accompany any such material before publication.

TITLE OF ARTICLE (hereinafter referred to as "the Work"):

COMPLETE LIST OF AUTHORS:

The undersigned hereby assigns to the IJECES all rights under copyright that may exist in and to the above Work, and any revised or expanded works submitted to the IJECES by the undersigned based on the Work. The undersigned hereby warrants that the Work is original and that he/she is the author of the complete Work and all incorporated parts of the Work. Otherwise he/she warrants that necessary permissions have been obtained for those parts of works originating from other authors or publishers.

Authors retain all proprietary rights in any process or procedure described in the Work. Authors may reproduce or authorize others to reproduce the Work or derivative works for the author's personal use or for company use, provided that the source and the IJECES copyright notice are indicated, the copies are not used in any way that implies IJECES endorsement of a product or service of any author, and the copies themselves are not offered for sale. In the case of a Work performed under a special government contract or grant, the IJECES recognizes that the government has royalty-free permission to reproduce all or portions of the Work, and to authorize others to do so, for official government purposes only, if the contract/grant so requires. For all uses not covered previously, authors must ask for permission from the IJECES to reproduce or authorize the reproduction of the Work or material extracted from the Work. Although authors are permitted to re-use all or portions of the Work in other works, this excludes granting third-party requests for reprinting, republishing, or other types of re-use. The IJECES must handle all such third-party requests. The IJECES distributes its publication by various means and media. It also abstracts and may translate its publications, and articles contained therein, for inclusion in various collections, databases and other publications. The IJECES publisher requires that the consent of the first-named author be sought as a condition to granting reprint or republication rights to others or for permitting use of a Work for promotion or marketing purposes. If you are employed and prepared the Work on a subject within the scope of your employment, the copyright in the Work belongs to your employer as a work-for-hire. In that case, the IJECES publisher assumes that when you sign this Form, you are authorized to do so by your employer and that your employer has consented to the transfer of copyright, to the representation and warranty of publication rights, and to all other terms and conditions of this Form. If such authorization and consent has not been given to you, an authorized representative of your employer should sign this Form as the Author.

Authors of IJECES journal articles and other material must ensure that their Work meets originality, authorship, author responsibilities and author misconduct requirements. It is the responsibility of the authors, not the IJECES publisher, to determine whether disclosure of their material requires the prior consent of other parties and, if so, to obtain it.

- The undersigned represents that he/she has the authority to make and execute this assignment.
- For jointly authored Works, all joint authors should sign, or one of the authors should sign as authorized agent for the others.
- The undersigned agrees to indemnify and hold harmless the IJECES publisher from any damage or expense that may arise in the event of a breach of any of the warranties set forth above.

_____ _____

Author/Authorized Agent **Date**

CONTACT
International Journal of Electrical and Computer Engineering Systems (IJECES)
Faculty of Electrical Engineering, Computer Science and Information Technology Osijek
Josip Juraj Strossmayer University of Osijek
Kneza Trpimira 2b
31000 Osijek, Croatia
Phone: +38531224600,
Fax: +38531224605,
e-mail: ijeces@ferit.hr