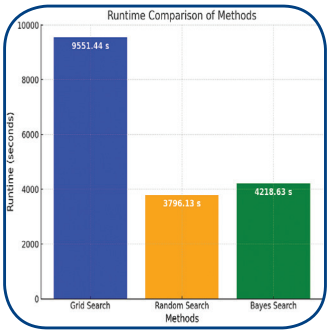
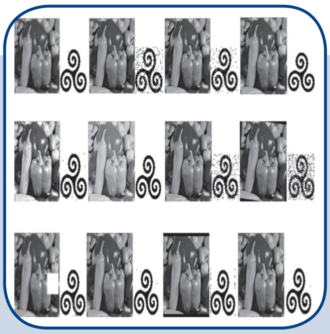
Volume 16, Number 6, 2025 ISSN 1847-6996

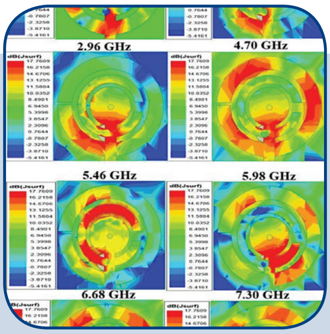
FERIT
FACULTY OF ELECTRICAL ENGINEERING, COMPUTER SCIENCE AND INFORMATION TECHNOLOGY **OSIJEK**

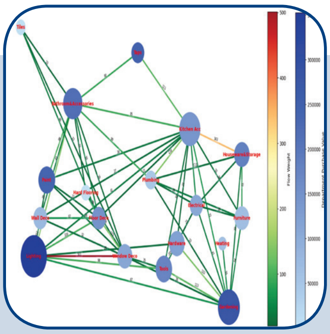
IJECES International Journal of Electrical and Computer Engineering Systems

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 16, Number 6, 2025 | Pages 431 - 495

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

Comparative Study of the Grid Current Harmonic Attenuation in a Photovoltaic Generator Due to the Influence of the Synchronization Strategy

Original Scientific Paper

N.F. Guerrero-Rodríguez*
Pontificia Universidad Católica Madre y Maestra
PUCMM, Engineering Sciences
Av. Abraham Lincoln Esq. Romulo Betancourt,
2748¸Santo Domingo, Dominican Republic,
nf.guerrero@ce.pucmm.edu.do

F.A. Ramírez-Rivera
Pontificia Universidad Católica Madre y Maestra
PUCMM, Engineering Sciences
Av. Abraham Lincoln Esq. Romulo Betancourt,
2748¸Santo Domingo, Dominican Republic,
franciscoramirez@pucmm.edu.do

R.O. Batista-Jorge
Pontificia Universidad Católica Madre y Maestra
PUCMM, Engineering Sciences
Av. Abraham Lincoln Esq. Romulo Betancourt,
2748¸Santo Domingo, Dominican Republic,
r.batista@ce.pucmm.edu.do

*Corresponding author

Julio A. Ferreira
Pontificia Universidad Católica Madre y Maestra
PUCMM, Engineering Sciences
Av. Abraham Lincoln Esq. Romulo Betancourt,
2748¸Santo Domingo, Dominican Republic,
julioferreira@pucmm.edu.do

Robert Mercado-Ravelo
Pontificia Universidad Católica Madre y Maestra
PUCMM, Engineering Sciences
Av. Abraham Lincoln Esq. Romulo Betancourt,
2748¸Santo Domingo, Dominican Republic,
rxmr0001@ce.pucmm.edu.do

Abstract – *The present paper studies the current harmonic attenuation for four synchronization strategies commonly used in inverters for photovoltaic power generation. In a model of a 6kWp photovoltaic generator, the four synchronization strategies are implemented in the inverter controller. Real-time simulations are performed, for this purpose, the models of the photovoltaic generator and the utility grid were embedded in an OPAL-RT OP5707XG simulator. The distortion of the grid current for each synchronization strategy is analyzed using its corresponding frequency spectrum and a comparison is made with the IEEE Std. 519-2022 standard. The purpose of this test is to observe the effect of the synchronization strategy on the harmonic attenuation of the grid currents. The results show that one of the synchronization strategies evaluated in this work may be sufficient for the system to comply with the harmonic standard for photovoltaic generators without the use of harmonic compensation structures or active harmonic filters. The evaluation of the harmonic distortion behavior of photovoltaic generators as a function of the synchronization strategy used in the inverter controller is the principal work contribution.*

Keywords: *Synchronization Strategy, Grid-Connected Photovoltaic Generator, Harmonic Distortion, Inverter, Real-Time Digital Simulator*

Received: July 11, 2024; Received in revised form: January 9, 2025; Accepted: January 10, 2025

1. INTRODUCTION

The large number of renewable energy resources (RES) that are integrated into the electrical power distribution system presents several issues, such as harmonic distortion in the utility grid [1-3]. Research has shown that the presence of diode rectifier-based systems and other nonlinear loads adversely affects the power quality of power systems, mainly due to the harmonic distortion generated by these types of systems [4-6].

Grid voltages are measured in grid-connected photovoltaic (PV) generators and sent to the control system of the inverters that act as power conditioners. These voltages are used in the control block as the synchronization strategy for determining grid parameters like grid phase angle and grid frequency, and to perform operations such as applying the Park transformations, a PQ power control, among other control techniques applied to these renewable agents. If the grid voltages are affected by harmonic distortions, these harmonics can be introduced into the inverter control system [7], so that the grid currents injected by the PV system contain harmonics coming from the grid voltages [8]. According to [7], if there is harmonic pollution in the grid voltage, the output current will also be affected by the harmonic.

Because the incidence of harmonics can influence the level of quality of the energy injected by photovoltaic generators [5, 9], and the fact that the grid currents must fulfil some normative regarding the harmonic distortion, where a 5% of the Total Harmonic Distortion of current (THD_I) it is recommended [10], several works investigate solutions to the problems due to the grid harmonic distortion caused by non-linear loads.

In [2] the authors examine the effect of grid voltage harmonics on the performance of rectified power converters by formulating the voltage and current harmonics of the converter. A method for finding the inverter current harmonics caused by the PWM signal as a dependence of the load profile is presented by the authors in [3]. In [11], an innovative design method for a Proportional Resonant (PR) controller with a selective harmonic cancellation is described for a grid-connected PV generator.

An integrated approach for improved grid current harmonic compensation including power ripple attenuation in a PV generator is proposed in [12]. In [13, 14], harmonic compensation structures are used to reduce current harmonic levels. Active harmonic filters are also used for this purpose, but there are no feasible for small grid-connected renewable energy system [5].

Previous works had focused on the effect and the techniques needed for harmonics compensation [6], arguments in favor of the need for the application of techniques for the reduction of the presence of harmonics in electrical power distribution systems with a strong presence of RES.

In addition, as can be seen in [15], the harmonics introduced by Inverter-Based Resources (IBR) under a high penetration of RES can, under certain circumstances, impact the system stability.

The ability of a PV inverter to attenuate harmonic currents due to the influenced of the synchronization strategy used in the inverters of photovoltaic generators is analyzed in [16], where the results showed that the synchronization strategy affects the harmonic attenuation capability of the inverter.

Due to the above, the impact of the synchronization strategy on reducing the harmonic distortion in grid-connected PV generators, is analyzed in this work to show their impact in the grid current harmonic attenuation of PV generators when voltage harmonics affect the utility grid.

To carry out the above mentioned, four synchronization strategies will be implemented in the MATLAB/SIMULINK R2022a software [17]. All of them will be simulated in an OPAL-RT's OP5707XG simulator [18] and their behavior will be compared to each other to have a clear analysis of their effect on the grid current harmonic attenuation. For this, the harmonic attenuation of the utility grid current due to the applied synchronization strategies will be analyzed using the corresponding frequency spectrum, and a comparison with the standard IEEE Std. 519-2022 [10] will be made.

The current work is organized in 6 sections: Section 2 provides a brief description of the synchronization strategies that will be evaluated in this work. Section 3 describes a PV generator which will be used as a case study to evaluate the impact that synchronization strategies have on the attenuation of harmonics in the grid currents. The methodology used in this work is shown in Section 4. In Section 5, simulations using MATLAB/SIMULINK R2022a software [17] and an OPAL-RT's OP5707XG simulator [18] are carried out. In Section 6 of this work, the results obtained are discussed and the conclusions are presented.

2. SYNCHRONIZATION STRATEGIES

To properly match the grid currents of the PV inverter to the grid voltages, it is needed to determine the phase angle of the grid [19-21]. A summary description of four synchronization strategies to be used in this comparative study is shown in Table 1.

Fig. 1a displays the diagram of the synchronization strategy named Phase-Locked Loop (PLL), which consists of the Clarke and Park transformations, a PI regulator and a voltage-controlled oscillator [22-24]. Fig. 1b shows the PSD+dqPLL synchronization strategy [22, 25], which consists of a PLL plus a Positive Sequence Detector (PSD) block able of determining the positive sequence of unbalanced grid voltages.

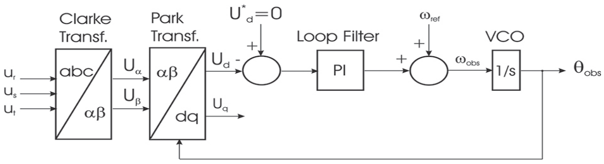
A Dual Second Order Generalized Integrator Frequency-Locked Loop (DSOGI-FLL) [24] synchronization strategy is shown in Fig. 1c, which consists of a block where the Clark transform is applied to the grid voltages, two blocks Quadrature Signal Generation (SOGI-QSG), a Positive Sequence Calculator (PSC) and a Frequency-Locked Loop (FLL).

Fig. 1d shows a Multiple Second Order Generalized Integrator Frequency-Locked Loop (MSOGI-FLL) synchronization strategy, which consists of a Harmonic Decoupling Network (HDN) block used to reject the influence of any other harmonics in the selected harmonic, four DSOGI-QSG blocks (for the 1st, 2nd, 5th and 7th harmonics), and a FLL block [26].

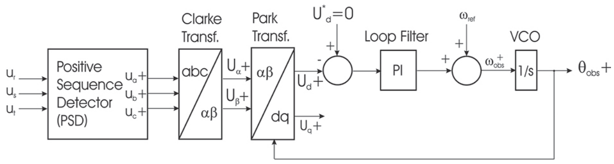
Since this work does not aim at a detailed study of each of these algorithms, Table 1 includes references of these algorithms, therefore that interested readers can study its design and operation in detail.

Table 1. Description of the synchronization strategies evaluated in this study

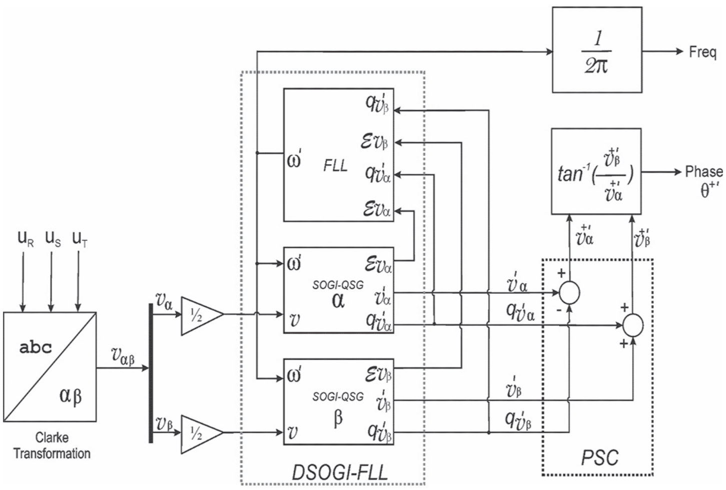
Reference	Name	Acronym	Locked methodology	Digital Filter used	Transfer function
[22-24]	Synchronous Reference Frame Phase-Locked Loop	dqPLL	Phase	Proportional Integrator (PI)	$H(S)=(Kps+Ki)/(s^2+Kps+Ki)$ (1)
[16, 22, 25]	Positive Sequence Detector plus a dqPLL	PSD+dqPLL	Phase	Proportional Integrator (PI)	$H(S)=(Kps+Ki)/(s^2+Kps+Ki)$ (1)
[24]	Dual Second Order Generalized Integrator Frequency-Locked Loop	DSOGI- FLL	Frequency	Second Order Generalized (SOGI)	$SOGI(s)=v'/(kE_v)(s)=\omega' s/(s^2+\omega'^2)$ (2)
[26]	Multiple Second Order Generalized Integrator Frequency-Locked Lop	MSOGI-FLL	Frequency	Second Order Generalized (SOGI)	$SOGI(s)=v'/(kE_v)(s)=\omega' s/(s^2+\omega'^2)$ (2)



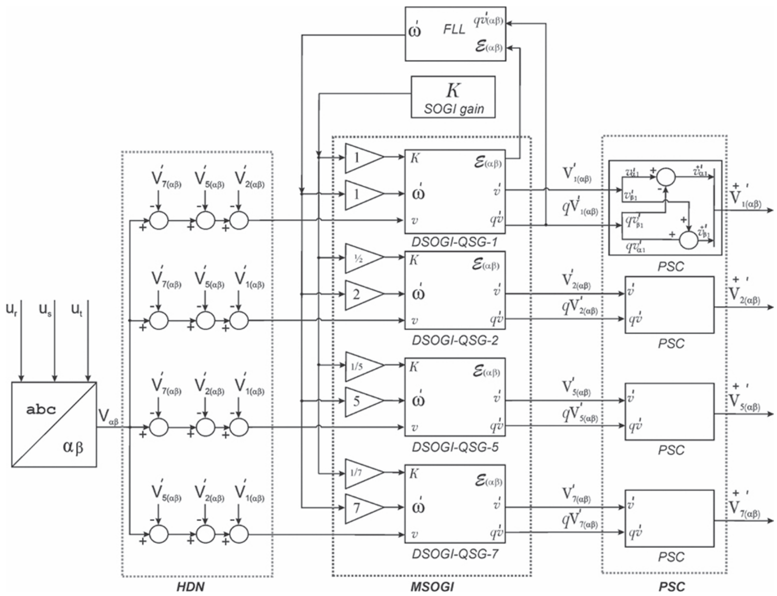
(a)



(b)



(c)



(d)

Fig. 1. Block diagrams of the synchronization strategies used in this study [27]. (**a**) dqPLL. (**b**) PSD+dqPLL. (**c**) DSOGI-FLL. (**d**) MSOGI-FLL

3. PHOTOVOLTAIC GENERATOR FOR CASE STUDY

A 6kWp PV 3-phase grid-connected generator model is used to evaluate the achieved grid current harmonic attenuation for each synchronization strategy analyzed, which is composed of two subsystems, Power and Control. The PV modules [28], the capacitor link, the inverter [29], the LCL filter [30], a transformer [31], an EMI filter [32], and finally, the utility grid [33] form the power subsystem. A maximum power point tracker (MPPT) block [28], [34], a cascade control [35] and the synchronization strategy form the control subsystem.

Table 2 summarized the Power subsystem parameters. The parameters used to design the synchronization strategies are shown in Table 3.

4. METHODOLOGY

Fig. 2 shows a flowchart of the methodology used in this work. A MATLAB/SIMULINK model of a 6kWp PV generator is developed using the parameters shown in Tables 2 and 3, implementing the synchronization strategies shown in Section 2. After validating the correct functioning of the MATLAB/SIMULINK model, the distortion caused by non-linear loads and converters used by renewable energy sources is simulated by voltage harmonics introduced into the 3-phase low-voltage utility grid. For this purpose, a 4% harmonic contamination in the 5th and 7th harmonics in the grid voltages is applied, attaining a Total Harmonic Distortion of voltage (THD_V) of 5.66%. For modelling the voltage grid with harmonic distortion, the system voltage of Eq. (3) is used.

$$
\begin{aligned}
U_{gr(t)} = {} & U_{1r}\cos(\omega t + \theta_{U1r}) \\
& + U_{h5r}\cos(5\omega t + \theta_{Uh5r}) \\
& + U_{h7r}\cos(7\omega t + \theta_{Uh7r}) \\
U_{gs(t)} = {} & U_{1s}\cos(\omega t + \theta_{U1s}) \\
& + U_{h5s}\cos(5\omega t + \theta_{Uh5s}) \\
& + U_{h7s}\cos(7\omega t + \theta_{Uh7s}) \\
U_{gt(t)} = {} & U_{1t}\cos(\omega t + \theta_{U1t}) \\
& + U_{h5t}\cos(5\omega t + \theta_{Uh5t}) \\
& + U_{h7t}\cos(7\omega t + \theta_{Uh7t})
\end{aligned} \tag{3}
$$

where h is the harmonic order, ω is the angular fundamental frequency, U is the voltage magnitude, θ is the voltage phase angle of the fundamental frequency, and θ_{Uh} is the voltage phase angle of the harmonic h.

Table 2. PV power subsystem parameters

Parameters	Value
Utility grid	$Vrms$=220V (ph- ph), 50Hz
Utility grid Total Harmonic Distortion of voltage	THD_V=5.66%
Outer Filter	R=20 mΩ, C=1.5 μF, L=3.9mH
Transformer	50Hz, 308/232V, $Snom$=6 kVA
Voltage Source Inverter	SKS22FB6U+E1CIF+B6CI, SEMISTACK-IGBT
DC bus voltage	Vcc= 350 V
PV power	P=6 kW

The synchronization strategies evaluated must be able to identify the frequency and phase of the 3-phase grid voltages in the presence of harmonic distortion, so that this information is sent to the PV generator control system to ensure proper inverter operation. If the phase and frequency of the grid are not detected during the MATLAB/SIMULINK simulations, it will be necessary to readjust the parameters governing the response and stability of the synchronization strategies studied, which are shown in Table 3.

After evaluating the operation of the photovoltaic generator in the presence of voltage harmonics, experiments will be performed using a real-time digital simulator. This allows tests to be carried out with certain similarities to those of a physical photovoltaic system.

The harmonic distortion of the utility grid current for each synchronization strategy is analyzed using its corresponding frequency spectrum and a comparison is made with the standard IEEE Std. 519-2022.

Table 3. Parameters of the synchronization strategies

Parameters	Value
PLL natural angular frequency	130 rad/s
Settling time	50 ms
Damping factor for the dqPLL	√2/2
Gain of the SOGI	√2
Damping factor for the DSOGI	0.707
Centre angular frequency for the DSOGI	314 rad/s
Gain to the settling time of the FLL block	100

The above is done to observe the impact of the synchronization strategies into harmonic attenuation of the currents. A calculation of the THD_I of the inverter grid currents is carried out for each synchronization strategy used, then the results obtained are compared and discussed.

The THD_I of the grid current is determined using the synchronization strategies shown in Table 1. The THD_I can be calculated using Eq. (4) [36].

$$
THD_I\,(\%) = \frac{\sqrt{[(I_2)^2 + (I_3)^2 + (I_4)^2 + (I_5)^2 + (I_n)^2]}}{I_1} X\ 100 \tag{4}
$$

where I_n and I_1 are the individual harmonic current distortion values in amperes and the fundamental current distortion values in amperes, respectively.

5. RESULTS

5.1. MATLAB SIMULATIONS

Modeling is done in MATLAB/SIMULINK2022a software [17] using the synchronization strategies shown in Table 1 and the values from Table 2 and Table 3. As a result of the MATLAB/SIMULINK simulations, Fig. 3a shows the grid voltages disturbed by a THD_V of 5.66%. The deformation of the 3-phase grid voltages due to the

5th and 7th harmonics can be observed. The distortion is intentionally added to the utility grid so that when these harmonics are sensed for feedback to the PV generator control subsystem, they interfere with the control subsystem. When the grid voltage signals are sent, the ability to identify the frequency and phase of the grid could be affected. This will allow the effect of synchronization strategies to cancel current harmonics at the inverter output when the grid is affected by voltage harmonics.

Because of harmonic distortion in the grid voltages, Fig.3b-c show the detected angular frequencies and phases by the synchronization strategies.

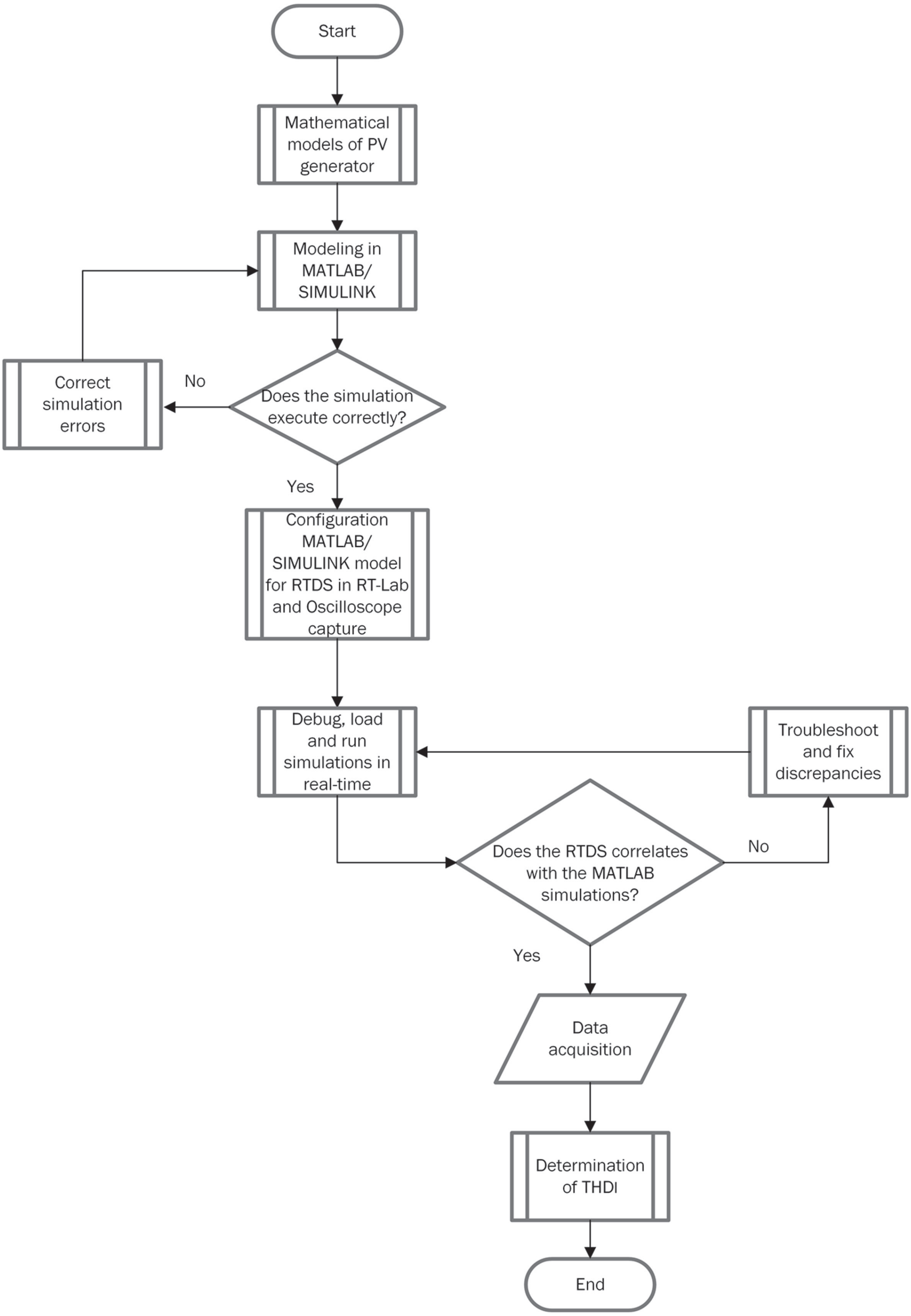


Fig. 2. Flowchart of the methodology to be used

Fig. 3b shows significant oscillations in the frequencies estimated by dqPLL and PSD+dqPLL (red and black, respectively), demonstrating the poor ability of these synchronization strategies to detect the grid frequency under conditions of harmonic distortion. In blue of Fig. 3b, a smaller oscillation is observed in the frequency detected by DSOGI-FLL, which allows a more approximate estimation of the grid frequency with respect to the frequencies obtained by dqPLL and PSD+dqPLL synchronization strategies. A better detection of the frequency is done by the MSOGI-FLL, as can be observed in yellow. This allows the control signals obtained from the synchronization strategy are not affected by the harmonic distortion of the grid voltages shown in Fig. 3a. As shown in Fig. 3c, poor phase detection is seen with dqPLL (red) and PSD+dqPLL (black); this is mainly due to the setting of the natural angular frequency of the PI regulator, which requires a trade-off between good harmonic attenuation and high dynamics. As shown in Fig.3c (yellow), near perfect phase detection is achieved with the MSOGI-FLL.

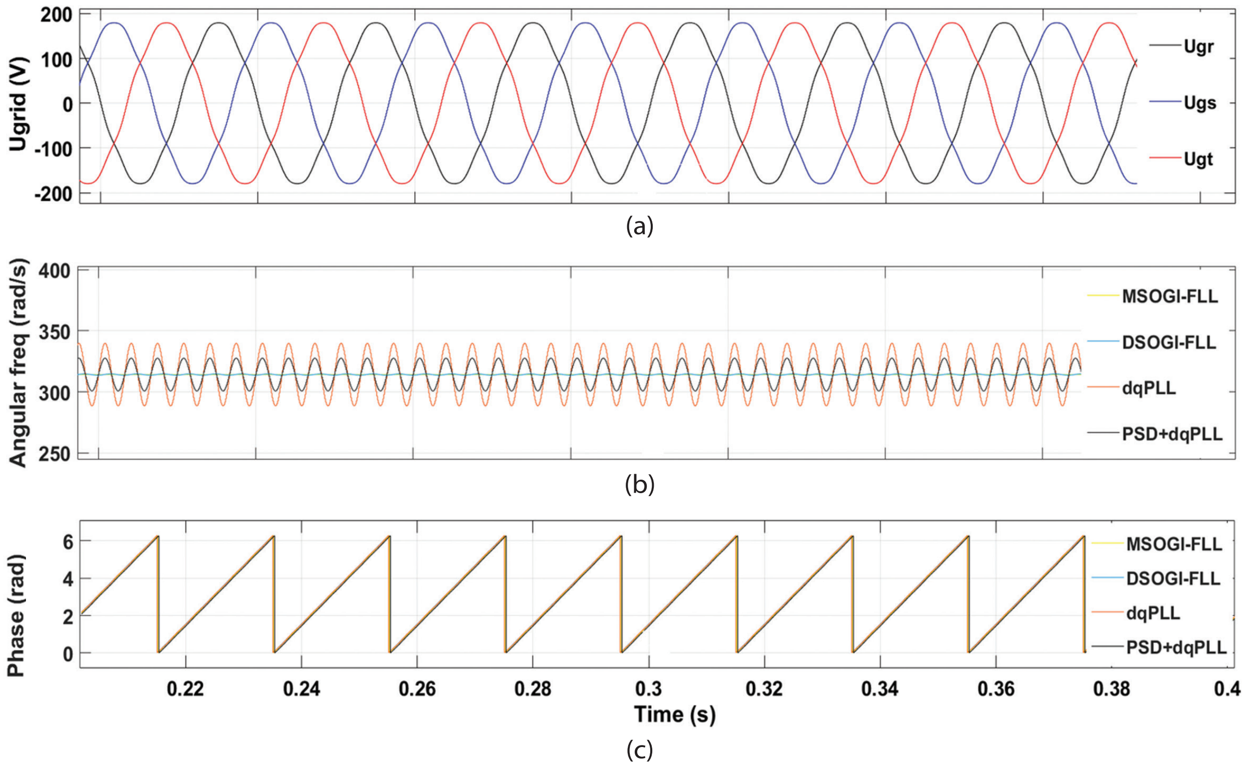


Fig. 3. (**a**) Grid voltages with harmonic pollution (THD_V = 5.66%.). (**b**) Time evolution of the detected angular frequency by the synchronization strategies under harmonic pollution. (**c**) Time evolution of the phase detection by the synchronization strategies under harmonic pollution

5.2. REAL-TIME DIGITAL SIMULATIONS

A photograph of the setup used to perform the real-time tests is shown in Fig. 4. It can be seen the OP5707XG real-time simulator, a server as the host PC and lastly, an oscilloscope is employed to record the voltage and current traces. A diagram of the real-time simulation configuration is displayed in Fig. 5 where can be seen the OP5707XG simulator, the Siglent SDS1204X-E oscilloscope and the PowerEdge R230 server computer host PC [37].

Fig. 6a shows the grid current when a dqPLL was used as the synchronization strategy, where there is significant distortion of the current waveform due to the presence of harmonics. As shown in the frequency spectrum in Fig. 6b, which corresponds to the grid current in Fig. 6a, the 5th and 7th harmonics are present. In percentage terms, the distortion at the 5th and 7th harmonics is 4.66% and 4.73%, respectively.

This results in a THD_I of 6.63%, which is higher than the maximum THD_I of 5% allowed by the IEEE Std. 519-2022 Standard.

The grid current when using a PSD+dqPLL synchronization strategy is shown in Fig. 7a. Comparing the current waveform with that obtained using dqPLL, a slight reduction in the harmonic distortion of the grid current can be observed. In the frequency spectrum in Fig. 7b, which corresponds to the grid current in Fig. 7a, the presence of the 5th and 7th order current harmonics can be observed, with percentages of 3.86% and 3.78% respectively, giving a THD_I=5.40%. However, despite the significant harmonic attenuation, the results do not comply with the limit proposed by the IEEE Std. 519-2022 Standard.

Fig. 8a shows the grid current using a DSOGI-FLL as a synchronization strategy. There is distortion of the current waveform due to the presence of harmonics. As can be seen from the frequency spectrum in Fig. 8b, which corresponds to the grid current in Fig. 8a, the 5th and 7th harmonics are present. In percentage terms, the distortion at the 5th and 7th harmonics is 3.69% and 3.67%, respectively. A THD_I=5.20% of the grid current was achieved, this value being higher than the 5% suggested by the regulations. Regarding the results obtained with the dqPLL and PSD+dqPLL synchronization strategies, significant harmonic attenuation was achieved when DSOGI-FLL was used.

Fig. 9a shows the grid current using a MSOGI-FLL as a synchronization strategy. Note the distortion of the current waveform due to harmonic pollution, however, when comparing the waveform with the current waveforms using the other synchronization strategies (Fig. 6a, 7a and 8a), the current waveform distortion using MSOGI-FLL is lower. The 5th and 7th harmonics are present in the frequency spectrum in Fig. 9b, which corresponds to the grid current in Fig. 9a. Expressed as a percentage, the 5th and 7th harmonic distortions are 3.15% and 3.36% respectively.

The percentage values for the 5th and 7th harmonics reveal that a significant attenuation of harmonics was achieved, resulting in a THD_I in the grid current of 4.60%, in compliance with the IEEE Std. 519-2022 Standard.

Fig. 4. Photo of the set-up for the execution of the simulations in real-time

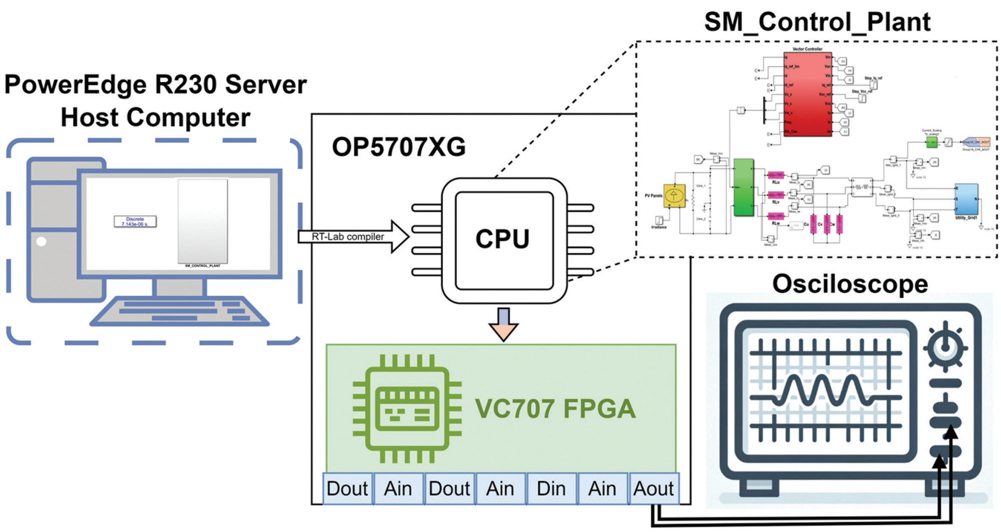
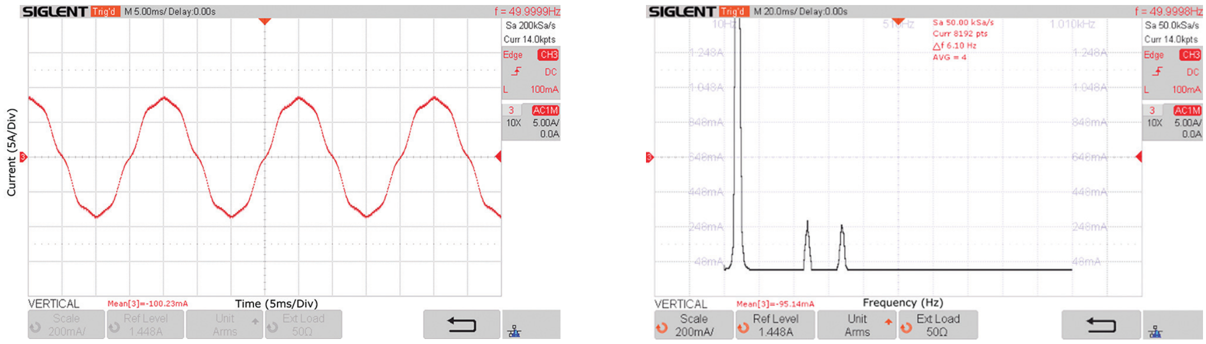


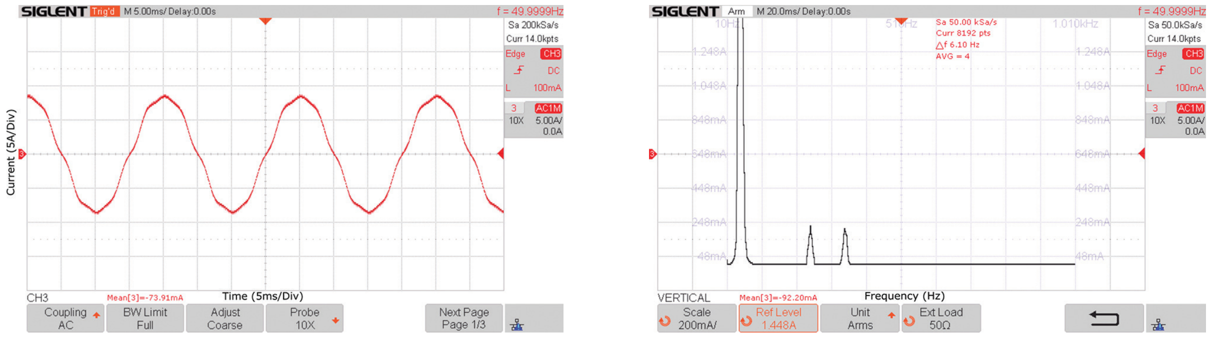
Fig. 5. Block diagram for the configuration of the simulations in real-time

The equivalent amplitude distortion in the grid current at phase 1 for the harmonics is shown in Table 4. Additionally, the resulting harmonic distortion when each synchronization strategy is applied is shown and compared to the IEEE Std. 519-2022 Standard.



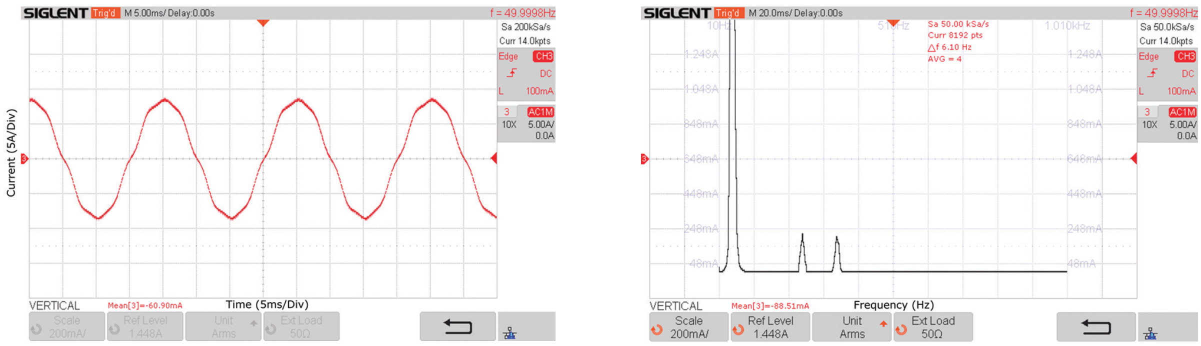
(a) (b)

Fig. 6. (**a**) Grid current corresponding to phase 1 when using a dqPLL is used. (**b**) Frequency spectrum corresponding to the grid current of phase 1 when a dqPLL is used



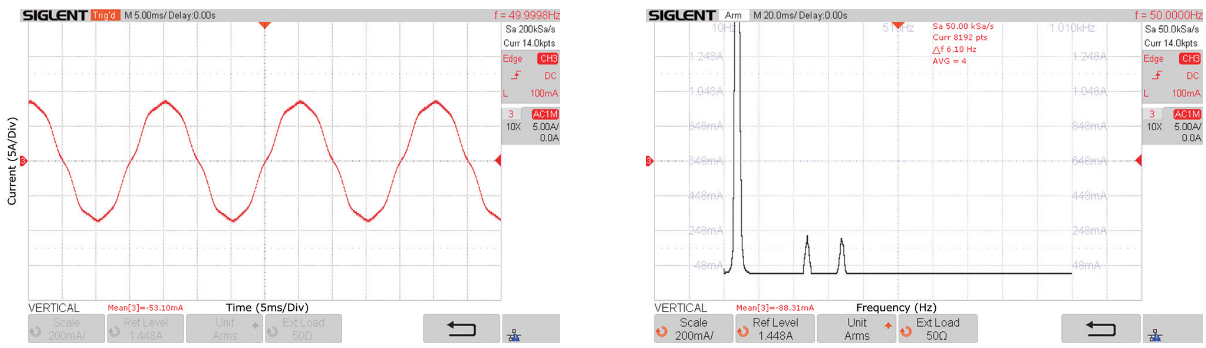
(a) (b)

Fig. 7. (**a**) Grid current corresponding to phase 1 when a PSD+dqPLL is used. (**b**) Frequency spectrum corresponding to the grid current of phase 1 when a PSD+dqPLL is used



(a) (b)

Fig. 8. (**a**) Grid current corresponding to phase 1 when a DSOGI-FLL is used. (**b**) Frequency spectrum corresponding to the grid current of phase 1 when a DSOGI-FLL is used



(a) (b)

Fig. 9. (**a**) Grid current corresponding to phase 1 when a MSOGI-FLL is used. (**b**) Frequency spectrum corresponding to the grid current of phase 1 when a MSOGI-FLL is used

Table 4. Harmonic distortion corresponding to the grid current of phase 1

Name	5th Harmonic distortion (%)	7th Harmonic distortion (%)	Distortion limit (%)	THD_I (%)	THD_I limit (%)
dqPLL	4.66	4.73	< 4.0	6.63	< 5.0
PSD+dqPLL	3.86	3.78	< 4.0	5.40	< 5.0
DSOGI-FLL	3.69	3.67	< 4.0	5.20	< 5.0
MSOGI-FLL	3.15	3.36	< 4.0	4.60	< 5.0

6. DISCUSSION AND CONCLUSION

Table 4 shows a summary of the THD_I of the grid current resulting from using the synchronization strategies shown in Table 1. Furthermore, these resulting THD_I values are evaluated with the IEEE Std. 519-2022 Standard. On the one hand, significant harmonic attenuation is observed when the MSOGI-FLL synchronization strategy is used. On the other hand, a poor harmonic attenuation is obtained when the dqPLL synchronization strategy was employed.

The contribution of four synchronization strategies to decrease the magnitude of current harmonics in PV generators when the grid voltage is disturbed by harmonic pollution was studied in this paper. Following, a resume of the performances of these synchronization strategies are remarked:

The dqPLL synchronization strategy uses a PI which behaves as a low pass filter. However, for the settling time used along the paper, the cutoff frequency is not enough to attenuate properly the harmonics, and, consequently, a poor harmonic attenuation is attained.

Compared to dqPLL, the notable attenuation of harmonic distortion achieved when using the PSD+dqPLL is due to the use of the positive sequence detector (PSD) block which cancels the effect of the negative sequence of the 5th harmonic. The above allows better detection of the grid phase and grid frequency, thus sending reliable information to the inverter controller used in the PV generator.

When the DSOGI-FLL is compared with the dqPLL, a greater reduction of harmonic contamination was obtained using the DSOGI-FLL. This better behavior is because a Positive Sequence Calculator block is used in the DSOGI-FLL. Despite the harmonic attenuation achieved, a significant harmonic contamination can be observed in the grid current, and then, the obtained THD_I exceeds the limits imposed by the standard on grid current harmonics IEEE Std. 519-2022.

When the MSOGI-FLL was tested, an important rejection in the harmonic distortion of the grid current harmonic was observed due to the use of the block Harmonic Decoupled Network [26], in addition to a DSOGI block corresponding to each of the harmonics found in the utility grid. Also, several Positive Sequence Calculator blocks were implemented, to attain only the positive sequence of the harmonic of the grid voltages and, as consequence, rejecting the harmonic pollution effects of the negative sequence.

This work concludes that depending on the synchronization strategy used, a different behavior of harmonic distortion attenuation is observed in the grid currents of photovoltaic generators. MSOGI-FLL is an effective solution to reject the magnitude certain number of harmonics.

ACKNOWLEDGMENTS:

This research was funded by the MESCyT (Ministry of Higher Education Science and Technology) in the Dominican Republic through Fondocyt, under the projects Development of New Methodologies for Harmonic Mitigation in Grid-connected Electrical Power Systems (2020-2021-3A9-068) and, Development of Methodologies Based on Solar-photovoltaic Green Hydrogen to Stabilize the Electrical Grid and Reduce the Carbon Footprint for Electrical Generation (2022-3C1-168).

7. REFERENCES:

[1] A. Alduraibi, “Analysis of harmonics generated by power converters in power networks: impacts and mitigation techniques”, School of Information Technology and Electrical Engineering, The University of Queensland, 2021, PhD Thesis.

[2] A. Alduraibi, J. Yaghoobi, F. Zare, “Impacts of Grid Voltage Harmonics Amplitude and Phase Angle Values on Power Converters in Distribution Networks”, IEEE Access, Vol. 9, 2021, pp. 92017-92029.

[3] J. Yaghoobi, F. Zare, “Impacts of three-phase power converter operating modes on harmonic emissions in distribution networks: Harmonics emission within 2-9 kHz”, IET Power Electronics, Vol. 13, No. 13, 2020, pp. 2935-2942.

[4] J. Yaghoobi, A. Alduraibi, D. Martin, F. Zare, D. Eghbal, R. Memisevic, “Impact of high-frequency harmonics (0-9 kHz) generated by grid-connected inverters on distribution transformers”, International Journal of Electrical Power and Energy Systems, Vol. 122, 2020, p. 106177.

[5] R. Ekström, M. Leijon, “Lower order grid current harmonics for a voltage-source inverter connected to a distorted grid”, Electric Power Systems Research, Vol. 106, 2014, pp. 226-231.

[6] A. Kalair, N. Abas, A. R. Kalair, Z. Saleem, N. Khan, "Review of harmonic analysis, modeling and mitigation techniques", Renewable and Sustainable Energy Reviews, Vol. 78, 2017, pp. 1152-1187.

[7] Y. Du, D. D. C. Lu, G. James, D. J. Cornforth, "Modeling and analysis of current harmonic distortion from grid connected PV inverters under different operating conditions", Solar Energy, Vol. 94, 2013, pp. 182-194.

[8] N. F. Guerrero-Rodríguez, A. B. Rey-Boué, "Modelling, simulation and experimental verification for renewable agents connected to a distorted utility grid using a Real-Time Digital Simulation Platform", Energy Convers Manag, Vol. 84, 2014, pp. 108-121.

[9] J. Badugu, Y. P. Obulesu, C. S. Babu, "Investigation of THD Analysis in Residential Distribution Systems with Different Penetration Levels of Electric Vehicles", International Journal of Electrical and Computer Engineering Systems, Vol. 9, No. 1, 2018, pp. 21-27.

[10] Institute of Electrical and Electronics Engineers IEEE, "519-2022 - IEEE Standard for Harmonic Control in Electric Power Systems", 2022.

[11] C. Yanarates, Z. Zhou, "Symmetrical Pole Placement Method-Based Unity Proportional Gain Resonant and Gain Scheduled Proportional (PR-P) Controller with Harmonic Compensator for Single Phase Grid-Connected PV Inverters", IEEE Access, Vol. 9, 2021, pp. 93165-93181.

[12] M. Mishra, V. N. Lal, "An Advanced Proportional Multiresonant Controller for Enhanced Harmonic Compensation with Power Ripple Mitigation of Grid-Integrated PV Systems under Distorted Grid Voltage Conditions", IEEE Transactions on Industry Applications, Vol. 57, No. 5, 2021, pp. 5318-5331.

[13] L. S. Xavier, A. F. Cupertino, J. T. de Resende, V. F. Mendes, H. A. Pereira, "Adaptive current control strategy for harmonic compensation in single-phase solar inverters", Electric Power Systems Research, Vol. 142, 2017, pp. 84-95.

[14] D. Zammit, C. S. Staines, M. Apap, J. Licari, "Design of PR current control with selective harmonic compensators using Matlab", Journal of Electrical Systems and Information Technology, Vol. 4, No. 3, 2017, pp. 347-358.

[15] Z. Deng, G. Todeschini, "A Novel Approach for Harmonic Assessment of Power Systems With Large Penetration of IBRs - A U.K. Case Study", IEEE Transactions on Power Delivery, Vol. 39, No. 1, 2024, pp. 455-466.

[16] N. F. Guerrero-Rodríguez, L. C. Herrero-de Lucas, S. de Pablo-Gómez, A. B. Rey-Boué, "Performance study of a synchronization algorithm for a 3-phase photovoltaic grid-connected system under harmonic distortions and unbalances", Electric Power Systems Research, Vol. 116, 2014, pp. 252-265.

[17] MathWorks, "MATLAB/SIMULINK", https://www.mathworks.com/products/matlab.html (accessed: 2024)

[18] OPAL-RT, "POWER HARDWARE-IN-THE -LOOP", https://www.opal-rt.com/power-hardware-in-the-loop/ (accessed: 2024)

[19] W. A. Sotelo, D. J. Rincon, M. A. Mantilla, J. M. Rey, Z. Zheng, "Performance Analysis of Three-Phase Synchronization Algorithms Under Voltage Sags", Proceedings of IECON 2022 - 48th Annual Conference of the IEEE Industrial Electronics Society, Brussels, Belgium, 17-20 October 2022, pp. 1-6.

[20] F. Blaabjerg, R. Teodorescu, M. Liserre, A. V Timbus, "Overview of control and grid synchronization for distributed power generation systems", IEEE Transactions on Industrial Electronics, Vol. 53, No. 5, 2006, pp. 1398-1409.

[21] S. V. Kulkarni, D. N. Gaonkar, "An investigation of PLL synchronization techniques for distributed generation sources in the grid-connected mode of operation", Electric Power Systems Research, Vol. 223, 2023, p. 109535.

[22] M. Karimi-Ghartemani, M. R. Iravani, "A Method for Synchronization of Power Electronic Converters in Polluted and Variable-Frequency Environments", IEEE Transactions on Power Systems, Vol. 19, No. 3, 2004, pp. 1263-1270.

[23] R. Teodorescu, M. Liserre, P. Rodriguez, "Grid Converters for Photovoltaic and Wind Power Systems", John Wiley & Sons, 2011.

[24] P. Rodríguez, R. Teodorescu, I. Candela, A. V. Timbus, M. Liserre, F. Blaabjerg, "New positive-sequence voltage detector for grid synchronization

of power converters under faulty grid conditions", Proceedings of the 37th IEEE Power Electronics Specialists Conference, Jeju, Korea, 18-22 June 2006, pp. 1-7.

[25] C. L. Fortescue, "Method of Symmetrical Co-Ordinates Applied to the Solution of Polyphase Networks", Transactions of the American Institute of Electrical Engineers, Vol. XXXVII, No. 2, 1918, pp. 1027-1140.

[26] P. Rodriguez, A. Luna, I. Candela, R. Mujal, R. Teodorescu, F. Blaabjerg, "Multiresonant Frequency-Locked Loop for Grid Synchronization of Power Converters Under Distorted Grid Conditions", IEEE Transactions on Industrial Electronics, Vol. 58, No. 1, 2011, pp. 127-138.

[27] N. F. Guerrero-Rodríguez, A. B. Rey-Boué, E. J. Bueno, O. Ortiz, E. Reyes-Archundia, "Synchronization algorithms for grid-connected renewable systems: Overview, tests and comparative analysis", Renewable and Sustainable Energy Reviews, Vol. 75, 2017, pp. 629-643.

[28] P. V. Mahesh, S. Meyyappan, R. Alla, "Support Vector Regression Machine Learning based Maximum Power Point Tracking for Solar Photovoltaic systems", International Journal of Electrical and Computer Engineering Systems, Vol. 14, No. 1, 2023, pp. 100-108.

[29] S. A. Zulkifli, E. Y. Riyandra, S. Salimin, A. N. Zaidan, R. Jackson, "Investigation of Multilevel Inverter for the Next Distributed Generation Using Low-Cost Microcontroller", International Journal of Electrical and Computer Engineering Systems, Vol. 10, No. 1, 2019, pp. 11-18.

[30] M. T. Faiz et al. "An alternative control synthesis for stability enhancement of digital-controlled LCL-filtered grid-connected inverter", Electric Power Systems Research, Vol. 213, 2022, p. 108678.

[31] L. Zhang et al. "Collaborative optimization control of hybrid distribution transformers integrated with photovoltaics in unbalanced distribution networks", International Journal of Electrical Power & Energy Systems, Vol. 159, 2024, p. 110052.

[32] L. Sandrolini, A. Mariscotti, "Impact of short-time fourier transform parameters on the accuracy of EMI spectra estimates in the 2-150 kHz supraharmonic interval", Electric Power Systems Research, Vol. 195, 2021, p. 107130.

[33] T. Y et al. "Optimization of Distributed Generation in Radial Distribution Network for Active Power Loss Minimization using Jellyfish Search Optimizer Algorithm", International Journal of Electrical and Computer Engineering Systems, Vol. 15, No. 3, 2024, pp. 215-223.

[34] M. Oubella, M. Ajaamoum, B. Bouachrine, "Low cost microcontroller implementation of Takagi-Sugeno Fuzzy MPPT controller for photovoltaic systems", International Journal of Electrical and Computer Engineering Systems, Vol. 13, No. 10, 2022, pp. 971-981.

[35] S. Buso, P. Mattavelli, "Digital Control in Power Electronics", Morgan & Claypool Publishers, 2006.

[36] M. Turarbek, A. Ruderman, B. Reznikov, "Calculation of Current Total Harmonic Distortion for a Single-Phase Current Source Multilevel Inverter", Proceedings of the IEEE 14th International Conference on Compatibility, Power Electronics and Power Engineering, Setubal, Portugal, 8-10 July 2020, pp. 249-254.

[37] N. F. Guerrero-Rodríguez, R. O. Batista-Jorge, F. A. Ramírez-Rivera, J. Ferreira, R. Mercado-Ravelo, A. Manilla, "Harmonic Distortion Study of a Photovoltaic Generator in a Microgrid under Disturbances", Energies, Vol. 17, No. 9, 2024, p. 2031.

Hyperparameter Optimization for Deep Learning Modeling in Short-Term Load Forecasting

Original Scientific Paper

Thanh Ngoc Tran
Faculty of Electrical Engineering Technology
Industrial University of Ho Chi Minh City,
12 Nguyen Van Bao Street, Ward 1, Go Vap District,
Ho Chi Minh City, Viet Nam
tranthanhngoc@iuh.edu.vn

Tuan Anh Nguyen*
Faculty of Electrical Engineering Technology
Industrial University of Ho Chi Minh City,
12 Nguyen Van Bao Street, Ward 1, Go Vap District,
Ho Chi Minh City, Viet Nam
nguyenanhtuan@iuh.edu.vn

*Corresponding author

Abstract *– The evolution of new technologies has made short-term power load forecasting an essential part of the streamlining process in the management of power grid systems. Machine learning algorithms have been applied widely in this area but with little success towards achieving better accuracy rates. These gaps point out the necessity for better forecasting methods . This study is about the power grid system from Ho Chi Minh city in Vietnam. Ho Chi Minh operates as a metropolitan area on the rise with economic activity and seasonal factors greatly influencing electricity consumption. Due to its intricate fluctuations in consumption pattern, the city is known for having a high level of energy. This makes the city suitable for an in-depth investigation regarding a case study on short-term load forecasting approaches. In this study, the goal is to evaluate the effectiveness of three hyperparameter optimization methods: Random Search, Grid Search, and Bayes Search. All these methods optimize the performance of Convolutional Neural Network (CNN) models for short-term electricity load forecasting in Ho Chi Minh City. The results obtained through this work can also be used as a basis for introducing the methods to other locations in Vietnam. The assessment of the techniques is performed using fundamental error measures such as Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE). Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), Mean Squared Error (MSE), and Root Mean Squared Error (RMSE). Bayes Search completed with an MAE of 77.93, MAPE of 2.94%, MSE of 10,376.7, and RMSE of 101.9. These results indicate a noticeable enhancement in prediction accuracy when compared with the outcomes from Grid Search and Random Search. Grid Search provided an MAE of 106.23, MAPE of 3.95%, MSE of 17,033.7, and RMSE of 130.5. Random Search produces results of an MAE of 96.8, MAPE of 3.57%, MSE of 14,951.0, RMSE of 122.3. These results are evidence that Bayes Search is better for short-term electricity load forecasting in Ho Chi Minh City. The study also proposes an evaluation framework, which is meant for load forecasting in Vietnam. It is designed for Ho Chi Minh City predicting purposes, thus, integrating innovative concepts with actual forecasting functions. The framework is also applicable to other areas in Vietnam, both rural and urban, having different power consumption patterns. The reduction in forecasting inaccuracies through the use of Bayes Search is found to be promising as observed in the research. This automation supports better decision-making in energy management. It helps reduce costs in dynamic and complex power grid environments. These findings have practical value. They support efforts to build more flexible and efficient energy grids in Vietnam.*

Keywords: *CNN network, hyperparameter optimization, Bayes Search, Random Search, Grid Search, Short-term load forecasting*

Received: September 30, 2024; Received in revised form: March 21, 2025; Accepted: April 7, 2025

1. INTRODUCTION

Load forecasting in the near future is essential for adequate power grid control in terms of power generation, distribution, and supply. Precise estimation of electricity consumption is key to preserving a stable provision of energy, and subsequently minimizing expenses.

It guarantees rational energy distribution, which is particularly important in major metropolitan regions. The demand for electricity in cities is volatile due to the expansion of the commercial and industrial activities. This necessitates the need for accurate forecasting. There is considerable improvement in forecasting techniques over the years. Still, there are difficulties.

One of the difficulties has always been the validity of the results. One issue is reliability. Conventional approaches, which bear this burden, have their own difficulties. Linear Regression [1], Time Series Forecasting [2], Kalman Filtering [3] are models that tend to fail for a variety of reasons. They have great difficulties deals with large, complex modern power system datasets. These datasets are often produced by contemporary power systems. To overcome these challenges, Deep Learning models are useful. Convolutional Neural Networks (CNN) is one of it. CNNs are capable of processing data with non-linear multi-dimensional structures. Thus, they are able to make better predictions [4]. They can uncover deeper patterns within the data than other methods. It also increases the overall accuracy of forecasts offered by the models.

Now we see more and more interest in load forecasting models in developing countries. There are so many techniques for improving forecasting. Study [5] proposed a hybrid model, HHO-GCN-LSTM. This model uses a sophisticated combination of deep learning architectures and optimization methods. It aims towards achieving more precise load forecasting. Study [6] recommends the development of ensemble based techniques for short term load forecasting. With this approach, wavelet transform, Extreme Learning Machine and Partical Least Squares Regression (PLSR) are used. The improvement of forecast accuracy is achieved by these methods. They also minimize overfitting. This enhances the reliability of the model's predictions. Study [7] introduces a spatial-temporal forecasting framework using Graph Neural Networks, leveraging individual and aggregated load data to capture hidden dependencies among residential units, significantly enhancing prediction accuracy over conventional methods. These studies demonstrate significant progress in load forecasting in developing countries and contribute to shaping advanced technological trends in energy management, particularly in the context of increasingly scarce resources that must be utilized efficiently. However, recent studies also suggest that integrating signal decomposition techniques with deep learning models, as demonstrated in Article [8], can result in higher accuracy for short-term forecasting.

In recent years, the development of short-term power load forecasting models has achieved notable advancements, driven by the progression of deep learning technologies. Studies [9-12] have explored these innovative methods, particularly the combination of Convolutional Neural Networks (CNNs) and Long Short-Term Memory Networks (LSTMs). Research [9] developed a Pyramid-CNN model for customers with similar energy usage profiles, enhancing forecasting accuracy through cluster analysis. Meanwhile, studies [10] have focused on integrating CNN and LSTM for short-term power load forecasting, demonstrating superior accuracy compared to traditional models. Further studies, such as those in [11], propose convolutional multi-integration models and convolutional wavelet models to handle complex fluctuations in load data effectively. Research [12] further expands model capabilities by utilizing evolutionary methods and encoder-decoder combinations in power load forecasting. Article [13] highlights the utilization of technical indicators such as EMA and chaotic optimization algorithms for time series prediction, showcasing their potential for applications in power load forecasting. Building on this foundation, research [14] delves into a CNN- LSTM combined deep learning model, demonstrating its high efficiency in addressing dynamic load fluctuations within specific power systems. Together, these studies underscore the advancements in predictive modeling techniques for power systems.

The key issue to note in the aforementioned advanced methods is that many studies have not adequately addressed hyperparameter optimization, which is a critical factor that can significantly influence the performance and accuracy of the model. Failure to optimize hyperparameters may lead to suboptimal outcomes, even though deep learning models like CNNs and LSTMs have demonstrated their efficacy in short-term power load forecasting. Integrating hyperparameter optimization techniques is thus an essential step to enhance the performance of these models further. Feature selection and model simplifications are covered in reference [15]. Some optimizational strategies, like FPA, outperform their non-optimized counterparts. These strategies cut down the cost for computation. Such techniques can be implemented in applications that need near real-time processing like power load forecasting. This pertains to hyperparameter tuning of the CNN model [16]. The aim is to refine load prediction in metropolitan areas. It employs sophisticated optimization techniques [17]. These techniques are Grid Search, Random Search and Bayes Search [18]. They assist in pinpointing optimal values for hyperparameters to maximize precision in load forecasting. In this case, the MAE, MAPE, RMSE, and MSE error metrics are used to validate the model's accuracy.

What stands out its novelty is that it analyses the power grid of Ho Chi Minh City in Vietnam. The research focuses on certain anomalies of electricity usage in the area. These anomalies are correlated with socioeconomics and seasonal changes. This research work differs from the general research works on the CNNs and the optimization methods. It focuses on the particular issues of the region. One of them is the problem of excessive volatility of electric power consumption. Another issue is the existing attitude regionally power systems applied modern optimization techniques. The research investigates three approaches to hyperparameter optimization. These were the Grid Search, Random Search, and Bayes Search methods. Each of them is assessed concerning the degree to which they are able to enhance the CNN model. The purpose of the modeling is to satisfy the conditions of the power grid system of Ho Chi Minh City.

This work attempt solve the problem of neglected hyperparameter tuning on the city's power grid systems.

It illustrates the use of Bayes Search in combination with self-training algorithm and how it enhances a customer's computing performance. The results show that Bayes Search can lead to more precise predictions compared to other methods. The analysis was performed on data obtained from a specific part of a city. It measures the performance of the developed models using MAE, MAPE, RMSE, and MSE which are considered error metrics. Evaluations of this nature are intended to complement the findings. This work commences further research in this area. It demonstrates the application of contemporary methods in load forecasting at short time intervals. Other cities which have the same power difficulties may use the methods presented in this stud.

2. THEORETICAL BASIS

2.1. CNN MODEL

Deep Learning is a part of machine learning. It uses artificial neural networks to process data in layers. It helps the model understand complex patterns. Convolutional Neural Network (CNN) is one type of neural network. It is made to process data in multiple arrays. Images are a common example. CNN models are widely used in vision tasks. These tasks include face detection, image recognition, and object identification.

A CNN has many parts. Some of them are key components. The convolutional layer is one of them. It uses filters to extract features from the input data. The activation layer is another part. It adds non-linearity. This helps the model learn complex patterns. The pooling layer works with the two layers. It reduces data size but keeps important information. The fully connected layer comes at the end. It brings together all extracted features. It handles the final classification or prediction [19].

Convolutional Layer: A filter (or filters) is slid through the input image to create a featured map in a convolutional layer. Each filter is small (3x3 or 5x5) and is applied to the entire input image to create a new featured map. The mathematics of this process can be represented as follows: For the input image I and filter F, the characteristic map is calculated by the convolutional product:

$$S(i, j)=(F*I)(i, j)= \sum_m \sum_n F(m,n)I(i-m,j-n) \tag{1}$$

Where:

(i, j): is the location on the characteristic map.

$S(i, j)$: The output value at position (i,j) after applying the convolution operation.

$F*I$: The convolution operation between the filter F and the input image I.

$F(m, n)$: The value at position (m,n) in the filter (also called a kernel).

$I(i-m, j-n)$: The value at position $(i-m, j-n)$ in the input image I.

$\sum_m \sum_n$: Summation over all elements of the filter.

Activation Layer: After the filter is applied, the values on the characteristic map are passed through a non-linear trigger function, usually ReLU (Rectified Linear Unit). The ReLU function is defined as:

$$ReLU(x)=max(0,x) \tag{2}$$

Where:

x: The input value to the activation function (which can be the output from a previous layer in a neural network).

$ReLU(x)$: The output value of the ReLU activation function.

$max(0,x)$: The function returns the more excellent value between 0 and x.

Pooling Layer: Pooling typically uses max pooling or average pooling to reduce the spatial size of featured maps, highlight essential features, and reduce the number of parameters. Maximum compounding is defined as:

$$P(i, j)=max_{k,\, l \in window}\, I(i+k, j+l) \tag{3}$$

Where:

$P(i, j)$: The output value of the max pooling operation at position (i, j).

$max_{k,\, l \in window}$: The maximum value is selected from the specified window.

$I(i+k, j+l)$: The input characteristic map (or feature map), where $i+k, j+l$ represents the elements inside the pooling window.

Fully Connected Layer: Data from the fully connected layer is flattened and fed into one (or more) fully connected layer (s). Each neuron in this layer is connected to all the neurons in the previous layer, each with its own weight. The output of this class is:

$$y=W_x+b \tag{4}$$

Where

x: is the input from the previous layer.

W: i s the weighted matrix.

B: is the bias vector.

CNN's mathematical model is complex and requires optimization during training to learn effective weights and filters. Typically, this is done with the aid of back-propagation algorithms. If there is optimization involved, one method is Gradient Descent.

2.2. CNN NETWORK HYPERPARAMETERS

CNNs are mostly used for image processing. However, they can also be used for time series tasks like power load forecasting. Many hyperparameters are important for this type of work. These parameters help define the model's performance and relevance.

- Fewer filters make it harder for the model to detect features. More filters help the model learn complex patterns. Too many filters increase computation time. The number of filters must balance complexity and performance.

- Larger filter sizes increase the visible area during convolution. They help the model capture broader patterns. Smaller filters focus on fine details. The filter size affects how well the model detects time-based patterns.
- Stride controls how far the filter moves. A large stride reduces computations but may miss important data. A small stride captures more detail but increases cost.
- Padding defines how the model treats the input edges. 'Same' padding keeps the input and output size equal. 'Valid' padding reduces the output size. Padding affects how the model reads edge data.
- Activation functions add non-linearity. ReLU is the most common. It helps reduce the vanishing gradient and speeds up training.
- Pooling layers reduce the number of feature maps. Max pooling and average pooling are the most used. Pooling helps reduce overfitting and focuses on key features.
- The learning rate controls how fast weights update. A high learning rate speeds up training but may cause errors. A low learning rate trains slowly but more carefully.
- Batch size is the number of training samples processed at once. Small batches train slowly but learn fine patterns. Large batches train fast but may miss details.
- Epochs define how many times the model goes over the data. More epochs give more learning chances. Too many may cause overfitting.
- Regularization helps reduce overfitting. L1 and L2 add penalties to large weights. Dropout turns off some neurons during training. These methods help the model generalize better.

Choosing correct values for all parameters improves CNN performance on time series data. Hyperparameters also affect the CNN structure and real-world use.

3.2. HYPERPARAMETER OPTIMIZATION METHOD

Tuning is a key component when training the machine learning model because it fine tunes the performance of the model. It determines the optimal values for a number of adjustable predefined settings like learning rate, batch size, hidden layers, neurons per layer, etc. All of these must be configured prior to commencing the training thus rendering them immutable during the training process. Hyperparameters are distinct from model weights which are changed during training. Hyperparameters remain unchanged thereby making tuning a time-consuming process that requires enormous computational resources. The accuracy of the model on new data is improved, though. In the absence of tuning, overfitting or underfitting of the model is likely resulting in poor performance in real life scenarios.

Grid Search

Grid search is a method for hyperparameter optimization. It tests all possible combinations of hyperparameter values. It works well in small search spaces and when high precision is needed. Bergstra and Bengio [20] said Grid Search is easy to learn and use for beginners. However, it becomes inefficient in large search spaces because it takes a lot of time and resources. For example, three hyperparameters with four values each require 64 runs. Hutter et al. [21] called it "brute-force" and said it does not learn from past trials. Because of this, deep learning and decision tree models often use other methods. Random Search and Bayesian Search are common alternatives. They focus on promising regions and save resources. Petro and Pavlo Liashchynskyi [22] noted the strength of these methods.

Random Search

Among many methods, Random Search is a practical option for hyperparameter tuning. It works well in large and complex search spaces. Grid Search tests every possible configuration. In contrast, Random Search tests a limited number of random configurations. Navon and Bronstein [23] showed that Random Search can still give good results. It also uses fewer resources, which helps when computational power is limited. First, the user selects the hyperparameters. Then, the user sets the value ranges. The algorithm randomly generates a few configurations to test. Florea and Andonie [24] suggested a new version called Weighed Random Search. This version still uses random generation. However, it adds probabilistic rules to focus on better parts of the search space. This makes the search more efficient. Still, Random Search is not always precise. If too few configurations are tested, the results may be poor. In such cases, Bayes Search is better. Bischl et al. recommended it for better accuracy [25]. This method learns from past tests. It helps the model find the best areas to search.

Bayes Search

Like other forms of hyperparameter tuning, BayesSearchCV attempts to minimize some objective function. The difference is its use of Bayesian optimization, an advanced technique used to optimize functions that are costly or time-consuming to evaluate. It usually employs Gaussian Processes to try and direct the search to better places. The model first looks to find the promising regions within the space to search. After that, it tries to focus on evaluating those regions. Compared to Grid Search or Random Search, this approach considerably lowers the amount of attempts required [26]. The model decides upon a new set of parameters, evaluates the function and adjusts the estimate of its performance accordingly. BayesSearchCV is especially great when resources are limited [27]. However, it sacrifices some accuracy due to the surrogate model. It also needs extra settings which makes it harder to use. Still, it performs better with complex models. Deep neural networks, for example, benefit a lot. They need fewer trials to find the best configurations [25].

3. PROPOSED OPTIMIZATION MODEL STRUCTURE

3.2. ALGORITHMIC FLOWCHART

The authors explain an algorithm flowchart in this article. Its purpose is to evaluate the efficiency of three optimization methods which include: Grid Search, Random Search, and Bayes Search. The analysis is conducted on the CNN model. Figure 1 shows the flowchart.

Where:

X_{train}, Y_{train}: are the training inputs for the model.

X_{test}, Y_{test}: are the testing inputs for the model.

Y_{pred}: symbolizes the output from the model after execution.

$Y1$, $Y2$,…,Yn: represent the past load values

The main processes of the flowchart are described as follows:

- Input Data: The process begins with the input data, consisting of historical load values ($Y1$, $Y2$,…,Yn).
- Input Data Processing: The input data is preprocessed and divided into Training Data (X_{train}, Y_{train}) used to train the CNN model and testing Data (X_{test}, Y_{test}) used to evaluate the model's performance.
- Search Space of Hyperparameters: A predefined search space of hyperparameters is established, specifying possible configurations for the CNN model.
- Optimization Algorithms: Three optimization algorithms, Grid Search, Random Search, and Bayes Search, are applied to explore the hyperparameter search space and identify the Optimal Hyperparameters.
- Training the CNN Model: The CNN model is trained on the training dataset (X_{train}, Y_{train}) using the optimal hyperparameters.
- Prediction: The trained CNN model predicts the output (Y_{pred}) based on the test input data (X_{test}).
- Evaluation: The predicted values (Y_{pred}) are compared with the actual test data (Y_{test}), and the errors are calculated using metrics such as MAE, MAPE, MSE, and RMSE.

The following formula describes the mathematical model of the error rates used in the paper. Where y_i is the actual value and $\hat{y}_i$ is the predicted value.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \quad (5)$$

$$MAPE = \frac{100}{n}\sum_{i=1}^{n}\left|\frac{y_i-\hat{y}_i}{y_i}\right| \quad (6)$$

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \quad (7)$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \quad (8)$$

Where

MAE: Mean Absolute Error.

$MAPE$: Mean Absolute Percentage Error.

MSE: Mean Squared Error.

$RMSE$: Root Mean Squared Error.

n: The total number of data points.

yi: The actual value at index i.

$\hat{y}_i$:The predicted value at index i.

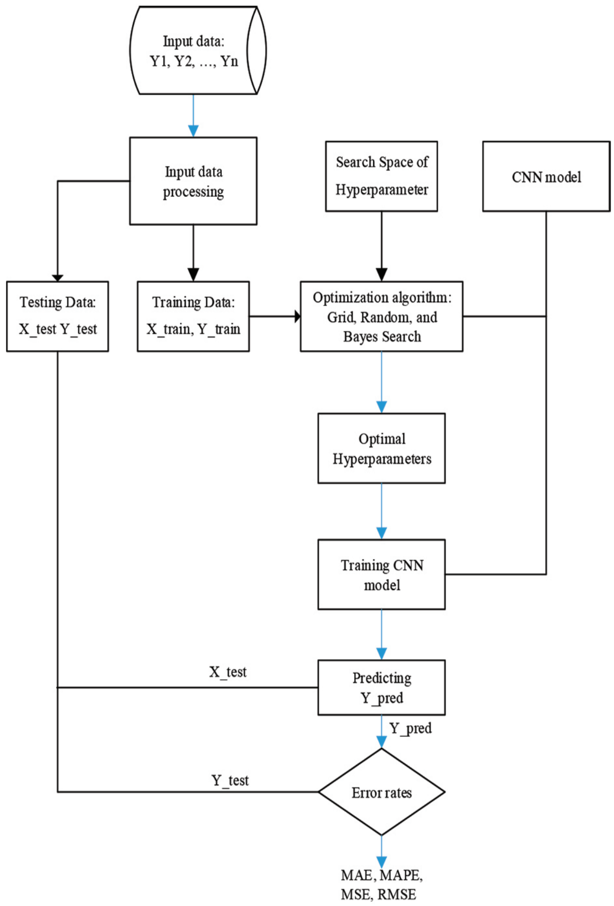


Fig. 1. Hyperparameter optimization algorithm diagram

4. CNN HYPERPARAMETER OPTIMIZATION ANALYSIS RESULTS

4.1. EXPERIMENTAL SETUP

In this study, the authors employed the electricity load dataset of Ho Chi Minh City, Vietnam, as presented in Table 1 below. The data sampling interval is 60 minutes, resulting in 24 data points daily. A sliding window approach with a window size of 24 generated Input-Target pairs (X, Y). The dataset (X, Y) consists of 840 samples, which were divided into a training dataset (X_{train}, Y_{train}) and a testing dataset (X_{test}, Y_{test}) with a ratio of 8:2.

Table 1. Historical Load Data in Ho Chi Minh City from 12/9/2016 to 31/12/2018

Date	00:00	01:00	……….	22:00	23:00
12/09/2016	1842.1	1795.1	……….	2337.2	2110.1
14/09/2016	1975.7	1914.6	……….	2297.5	2106.2
……….	……….	……….	……….	……….	……….
30/12/2018	2083.3	1980.9.	……….	2325.4	2127.8
31/12/2018	1902.7	1776.4	……….	2233.8	2059.5

Fig. 2 presents the electricity load profile for December 31, 2018. The chart shows an apparent fluctuation in the electricity load throughout the day, with the minimum load occurring in the early morning (1072.4 MW) and the maximum load during peak hours (2032.9 MW). This reflects the low electricity demand at night and early morning, while the demand increases significantly around midday when people and facilities use the most electricity. This chart helps identify usage trends throughout the day, allowing for better planning of an efficient power supply. Understanding these fluctuations not only aids in managing electricity distribution more effectively but also helps optimize operational strategies and distribution, ensuring that peak demand is met while saving resources during off-peak hours.

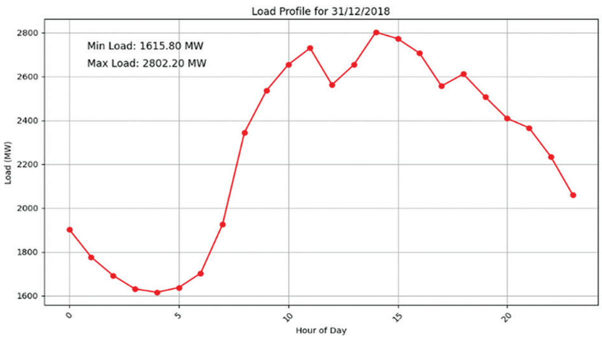


Fig. 2. Electricity Load on December 31, 2018

Table 2 presents the search space for the hyperparameters of the CNN model under investigation, including the number of filters, kernel size, batch size, and epochs. These search spaces are consistently applied across the Grid Search, Random Search, and Bayes Search algorithms.

Table 2. Hyperparameter search space

Hyperparameter	Search Space	Description
Filters	[16, 32, 64, 96, 128]	Number of filters in the Conv1D layer
Kernel Size	[3, 5, 7]	Size of the convolutional window
Batch Size	[16, 32, 64, 128, 256]	Number of samples processed in each training step
Epochs	[50, 100, 150, 200, 250]	Number of complete passes through the training dataset.

4.2. EXPERIMENTAL RESULTS

Table 3 presents the optimal hyperparameter values obtained from the three algorithms: Grid Search, Random Search, and Bayes Search. These results correspond to the hyperparameter search space described in Table 2 for the CNN model.

Table 3. Hyperparameter sets with optimal methods

	Randon Search	Gird Search	Bayes Search
Filters	32	32	64
Kernel Size	7	7	5
Batch Size	16	32	32
Epochs	250	250	200

Figs. 3, 4, and 5 illustrate how the predicted values from Grid Search, Random Search, and Bayes Search align with actual observations. The visual similarity between predicted and real values suggests that each algorithm successfully fine-tunes the CNN model to an acceptable level of accuracy. However, the differences in precision still matter when choosing the optimal method. The results prove the effectiveness and accuracy of the methodologies employed.

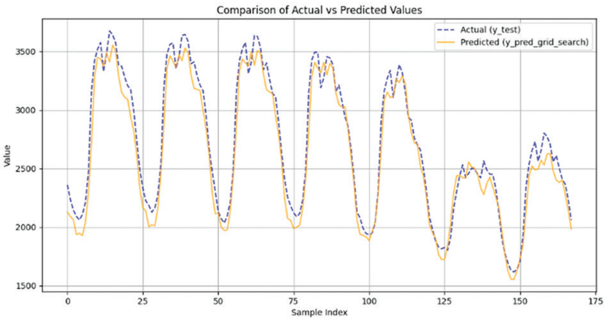


Fig. 3. The graph of predicted values using the Grid Search algorithm

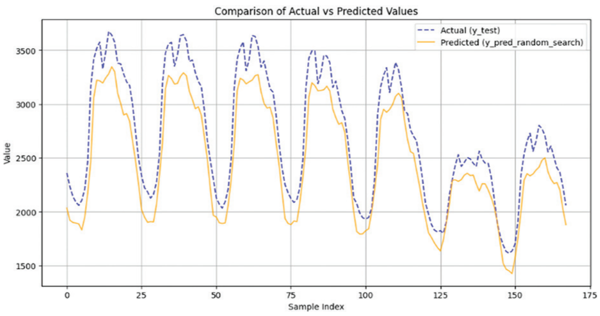


Fig. 4. Predicted values graph using Random Search

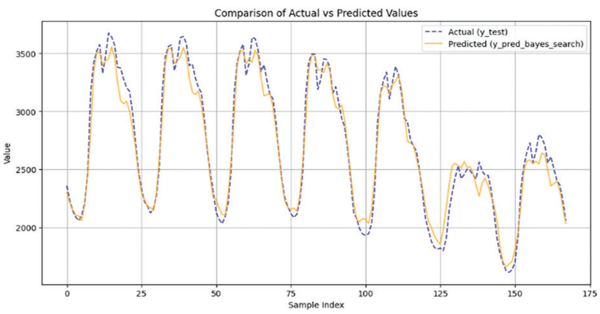


Fig. 5. Predicted values graph using Bayes Search

The error metrics for Grid Search, Random Search, and Bayes Search are shown in Table 4. The metrics include MAE, MAPE, MSE, and RMSE. Bayes Search gives the best results. It has the lowest MAE of 81.94. It also gives a MAPE of 3.09%, an MSE of 11,458.0, and an RMSE of 107.1. These values show high accuracy. They also show that Bayes Search improves model robustness. This makes it the most effective method. Random Search gives the worst results. It has an MAE of 166.7, a MAPE of 5.92%, an MSE of 36,783.19, and an RMSE of 191.7. These values show poor performance in finding good hyperparameters. Random Search is simple but not efficient here. Grid Search performs better than Random Search. It gives an MAE of 124.15, a MAPE of 4.58%, an MSE of 22,116.2, and an RMSE of 148.7. Grid Search can find near-optimal values. However, it is still less precise than Bayes Search.

Table 4. The results of error rates

Search_Method	MAE	MAPE	MSE	RMSE
Grid Search	124.15	4.58	22116.2	148.7
Random Search	166.7	5.92	36783.19	191.7
Bayes Search	81.94	3.09	11456.0	107.1

Fig. 6 presents the execution time for the three algorithms: Grid Search, Random Search, and Bayes Search. The runtime comparison shows that Grid Search is the slowest, taking nearly 9,551.44 seconds due to its exhaustive evaluation. Random Search improves efficiency with a runtime of around 3,796.13 seconds. Bayes Search is the fastest, completing it in just over 4,218.63 seconds, making it the most efficient method, according to the execution time.

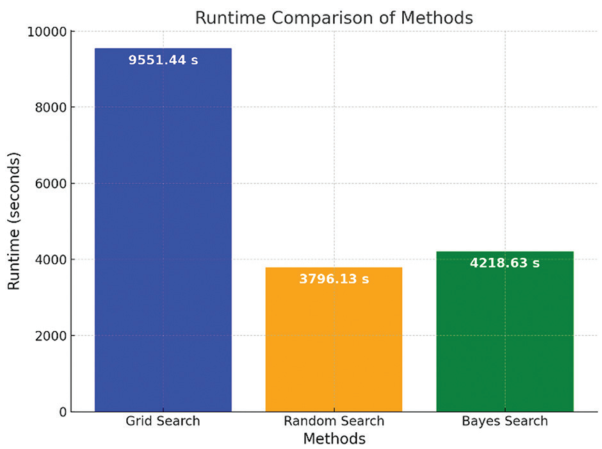


Fig. 6. Runtime of Optimization Methods

5. CONCLUSION

This study underlines how hyperparameter optimization—using techniques like Grid Search, Random Search, and Bayes Search—can significantly enhance CNN model performance for forecasting electricity demand. Among them, Bayes Search showed the highest predictive accuracy and the shortest processing time. In contrast, while Grid Search was moderately accurate, its longer runtime made it less practical. Random Search was quicker but less precise, making it a less reliable option for pinpointing the best parameter combinations Although moderately accurate, the Grid Search algorithm had the longest execution time, reflecting its inefficiency for large-scale problems. The Random Search algorithm showed better runtime efficiency than the Grid Search algorithm but produced higher error values, making it less dependable for optimal configurations. Future studies could further investigate advanced optimization algorithms and their application to larger datasets to enhance forecasting performance.

6. REFERENCES

[1] T. Zhao, J. Wang, Y. Zhang, "Day-Ahead Hierarchical Probabilistic Load Forecasting with Linear Quantile Regression and Empirical Copulas", IEEE Access, Vol. 7, 2019, pp. 80969-80979.

[2] P. Li et al. "Dynamic Similar Sub-Series Selection Method for Time Series Forecasting", IEEE Access, Vol. 6, 2018, pp. 32532-32542.

[3] S. Sharma, A. Majumdar, V. Elvira, E. Chouzenoux, "Blind Kalman Filtering for Short-Term Load Forecasting", IEEE Transactions on Power Systems, Vol. 35, No. 6, 2020, pp. 4916-4919.

[4] Z. Deng, B. Wang, Y. Xu, T. Xu, C. Liu, Z. Zhu, "Multi-scale convolutional neural network with time-cognition for multi-step short-Term load forecasting", IEEE Access, Vol. 7, 2019, pp. 88058-88071.

[5] D. N. Hung, N. M. Tam, N. T. Linh,

N. T. Hoan, N. T. Duy "Application of combining data preprocessing with wavelet filtering for GCN-LSTM network with HHO optimization algorithm in load forecasting model", TNU Journal of Science and Technology, Vol. 229, No. 6, 2024, pp. 160-169.

[6] S. Li, L. Goel, P. Wang, "An ensemble approach for short-term load forecasting by extreme learning machine", Applied Energy, Vol. 170, 2016, pp. 22-29.

[7] W. Lin, D. Wu, B. Boulet, "Spatialoral Residential Short-Term Load Forecasting via Graph Neural Networks", IEEE Transactions on Smart Grid, Vol. 12, No. 6, 2021, pp. 5373-5384.

[8] C. Barış, C. Yanarateş, A. Altan, "A robust chaos-inspired artificial intelligence model for dealing with nonlinear dynamics in wind speed forecasting", PeerJ - Computer Science, Vol. 10, 2024, pp. 1-40.

[9] K. Aurangzeb, M. Alhussein, K. Javaid, S. I. Haider, "A Pyramid-CNN based deep learning model for power load forecasting of similar-profile energy customers based on clustering", IEEE Access, Vol. 9, 2021, pp. 14992-15003.

[10] S. H. Rafi, N. Al-Masood, S. R. Deeba, E. Hossain, "A short-term load forecasting method using integrated CNN and LSTM network", IEEE Access, Vol. 9, 2021, pp. 32436-32448.

[11] H. U. I. H. Goh, S. Member, B. He, "Multi-Convolution Feature Extraction and Recurrent Neural Network Dependent Model for Short-Term Load Forecasting", IEEE Access, Vol. 9, 2021, pp. 118528-118540.

[12] S. M. J. Jalali et al. "A Novel Evolutionary-Based Deep Convolutional Neural Network Model for Intelligent Load Forecasting", IEEE Transactions on Industrial Informatics, Vol. 17, No. 12, 2021, pp. 8243-8253.

[13] A. Altan, S. Karasu, "Crude oil time series prediction model based on LSTM network with chaotic Henry gas solubility optimization", Energy, Vol. 242, p. 122964.

[14] F. Xu, G. Weng, Q. Ye, Q. Xia, "Research on Load Forecasting Based on CNN-LSTM Hybrid Deep Learning Model", Proceedings of the IEEE 5th International Conference on Electronics Technology, Chengdu, China, 13-16 May 2022, pp. 1332-1336.

[15] İ. Yağ, A. Altan, "Artificial Intelligence-Based Robust Hybrid Algorithm Design and Implementation for Real-Time Detection of Plant Diseases in Agricultural Environments", Biology, Vol. 11, No. 12, 2022.

[16] H. Aldosari, "A Hybrid Support Vector Machine with Grasshopper Optimization Algorithm based Feature Selection for Load Forecasting", Proceedings of the International Conference on Distributed Computing and Optimization Techniques, Bengaluru, India, 15-16 March 2024, pp. 1-4.

[17] L. Yang, A. Shami, "On hyperparameter optimization of machine learning algorithms: Theory and practice", Neurocomputing, Vol. 415, 2020, pp. 295-316.

[18] M. Munem, T. M. Rubaith Bashar, M. H. Roni, M. Shahriar, T. B. Shawkat, H. Rahaman, "Electric power load forecasting based on multivariate LSTM neural network using Bayes Search", Proceedings of the IEEE Electric Power and Energy Conference, Edmonton, AB, Canada, 9-10 November 2020, pp. 5-10.

[19] R. Patel, S. Patel, "A comprehensive study of applying convolutional neural network for computer vision", International Journal of Advanced Science and Technology, Vol. 29, No. 6, 2020, pp. 2161-2174.

[20] J. Bergstra, Y. Bengio,"Random Search for Hyper-Parameter Optimization", Journal of Machine Learning Research, Vol. 13, 2012, pp. 281-305.

[21] C. A. Coello Coello, "Learning and Intelligent Optimization", Vol. 3618, Springer, 2005.

[22] P. Liashchynskyi, P. Liashchynskyi, "Grid Search, Random Search, Genetic Algorithm: A Big Comparison for NAS", arXiv:1912.06059, 2019.

[23] D. Navon, A. M. Bronstein, "Random Search Hyper-Parameter Tuning: Expected Improvement Estimation and the Corresponding Lower Bound", arXiv:2208.08170, 2022.

[24] A. C. Florea, R. Andonie, "Weighted Random Search for hyperparameter optimization", International Journal of Computers Communications & Control, Vol. 14, No. 2, 2019, pp. 154-169.

[25] B. Bischl et al. "Hyperparameter optimization: Foundations, algorithms, best practices, and open challenges", WIREs Data Mining and Knowledge Discovery, Vol. 13, No. 2, 2023, pp. 1-43.

[26] J. Snoek, H. Larochelle, R. P. Adams, "Practical Bayes Search of machine learning algorithms", Advances in Neural Information Processing Systems, Vol. 4, 2012, pp. 2951-2959.

[27] B. Shahriari, K. Swersky, Z. Wang, R. P. Adams, N. De Freitas, "Taking the human out of the loop: A review of Bayes Search", Proceedings of the IEEE, Vol. 104, No. 1, 2016, pp. 148-175.

Highly Miniaturized Octa-band Antenna Using Concentric Circular Split Ring Structures

Original Scientific Paper

Padmaja Priyadarsini P
School of Electronics Engineering
VIT-AP University
Amaravathi, Andhra Pradesh - 522237, India
priyadarsini.20phd7143@vitap.ac.in

Chandu DS *
School of Electronics Engineering
VIT-AP University
Amaravathi, Andhra Pradesh - 522237, India
chandu.ds@vitap.ac.in

*Corresponding author

Abstract *– The proposed antenna design represents a systematic evolution from a simple ring-shaped structure to a highly efficient and versatile multi-band configuration tailored for next-generation wireless communication systems. Through the integration of rectangular slots, circular-shaped split ring resonators (CSRRs), and optimized geometries, the antenna achieves octa-band operation with superior impedance matching, enhanced bandwidth, and improved performance over a wide frequency range. The antenna is designed to support communication bands, including 4G/3G/2G, Wi-Fi, WLAN, WiMAX, and sub-6 GHz 5G connectivity. It resonates at 1.22, 2.12, 3.00, 4.76, 5.40, 5.94, 6.70, and 7.44 GHz, with a compact electrical size of 0.08 λ0 × 0.08 λ0. The inclusion of CSRRs and slots significantly expands the operational frequency bands while achieving a radiation efficiency of 84% and a peak gain of 4.1 dBi, making it well-suited for modern wireless applications. Moreover, the antenna design simplifies manufacturing and reduces the costs. Its compact and efficient structure ensures seamless integration into portable devices such as tablets and laptops, addressing the challenges of system complexity, size, and cost. This antenna provides a practical and scalable solution to meet the demands of diverse frequency bands in modern communication systems.*

Keywords*:* *Antenna design, CSSRR Resonator, Octa band, Gain and radiation efficiency, HFSS*

Received: December 10, 2024; Received in revised form: April 8, 2025; Accepted: April 8, 2025

1. INTRODUCTION

Over the past two decades, the rapid evolution of modern wireless personal communication devices, such as smartphones, has driven the demand for antennas capable of operating across multiple frequency bands while maintaining omnidirectional radiation characteristics. The integration of various wireless standards [1], including Bluetooth, WLAN, IOT, 5G sub-6 GHz, Wi-MAX [2-4], and RFID, necessitates antennas that can support diverse frequency ranges within the spatial constraints of compact devices. This presents a significant challenge for antenna researchers in designing structures that balance miniaturization with performance.

To address these challenges, several antenna design techniques have been developed to support multi-band functionality. Multi-band antennas offer a distinct advantage by reducing system complexity and physical dimensions, as a single antenna can cater to multiple applications. Antenna design plays an important role in determining essential parameters such as bandwidth, radiation efficiency, and overall system performance.

Microstrip patch antennas have emerged as a preferred choice for most wireless applications due to their inherent advantages, including compact size, lightweight construction, low fabrication cost, planar structure, and the ability to operate across multiple frequency bands. These features make microstrip patch antennas suitable for integration into modern devices, eliminating the need for multiple antennas to accommodate different resonant frequencies [5-7]. Consequently, these antennas are extensively employed in contemporary wireless communication systems, aligning with the requirements of advanced technologies and applications.

Various topologies are used for enhancing multi-band antenna. These typologies require extra space in the system to accommodate the external multiplexer circuits needed to choose the frequency bands. Multiple bands can be used at once with the latest generation of antennas, which also have intrinsic multiplexing properties. Metamaterial structures commonly used include Split-Ring Resonators (SRRs) [8, 9] and Complementary Split-Ring Resonators (CSRRs) [10, 11]. SRRs typically consist of two split rings facing in opposite directions and can take on various shapes such as rings, circles, squares, triangles, and hexagons. The rings function as resonators because of the presence of inductors as well as capacitors in the metal ring, and the resonant frequency can be refined by altering aspects such as split gap, metal width, spacing between adjacent rings, adjacent length, as well as presenting multiple split gaps. Reference [6] showcases the design, fabrication, and testing of a hexa-band monopole antenna with a low profile. Applications for 3G Advanced, Wi-Fi, WLAN, and WiMAX are all within the scope of the intended planar monopole antenna. A microstrip patch antenna [10] is constructed and studied for six working bands with slight frequency ratios with slender RF channel frequencies. An innovative four inverted L-shaped slots on a patch antenna is suggested for use with hexa-band circular shapes presented in [14]. The suggested antenna, which has a small footprint and operates on the Hexa-band frequency, is well-suited for a variety of uses, including those involving RADAR, Bluetooth, 5G mid-band, WiMAX, WLAN, LTE, as well as Wi-Fi [15]. A miniaturized hexa- band monopole antenna with circular polarization is expected to have multiple uses in [16] it is possible that WLANs and other wireless communication equipment would benefit from the suggested antenna. Hexa-band, dual-polarization performance is demonstrated by the favored design [17] at frequencies of 3,46, 8.28, 12.26, 17.21, 23,40, and 26.01. The miniature hexagonal fractal antenna has excellent improvement, high directionality, and an omnidirectional radiation outline throughout the multi-resonant frequency for 5G, IoT, satellite, and radar applications. This design [18] features a hexa-band PBG stacked MSPA that resonates between 1 and 9 GHz. In [16], a hepta-band resonance antenna for fifth-generation (5G) technology is planned. It is modeled on a Taconic TLY-3 substrate and resonates at a variety of frequencies that have been defined, including 28.1 GHz, 36.7 GHz, 45.8 GHz, 55.2 GHz, 62.8 GHz, 72.3 GHz, and 82 GHz accordingly. Antennas detailed in references [19-26] may receive signals on eight different service bands: PCS, LTE700, GSM850, GSM900, UMTS, DCS, LTE2500, and LTE2300. However, they have not been able to receive signals on the Wi-Fi, WiMAX, WLAN, or 5G bands. A total of seven service bands DCS, PCS, GSM850, GSM900, UMTS, LTE2500, and LTE2300 are covered by the antennas mentioned in references [27-31]. However, the Wi-Fi, WLAN, WiMAX, and 5G bands are not among them. There are works that reported miniaturization [32, 33] but the antennas are not multi- band and the design is complex.

The study of existing literature reveals that wireless communication systems require compact antennas with excellent performance in terms of bandwidth, gain, and efficiency. This research primarily focuses on the design of compact antennas for multi-band wireless applications in modern technologies. Specifically, it explores the performance of a single antenna that supports multiple applications while maintaining a compact size and high performance.

Key Advantages of the proposed antenna:

- The proposed antenna features a compact design utilizing a circular-shaped split ring resonator (CSSRR) structure with a full ground plane.
- It achieves octa-band operation, making it suitable for modern wireless communication applications.
- The design demonstrates a peak gain of 4.71 dBi and maintains a VSWR below 2 across all operating bands.
- Radiation efficiency ranges from 77% to 85% over the entire operating frequency range.
- The antenna is well-suited for defense tracking and weather monitoring applications due to its unique features.
- It is excited using a 50-ohm impedance-matched transmission line with full ground as reflector.
- The compact dimensions are 20 mm × 20 mm, with an electrical size of 0.08 $\lambda 0 \times 0.08\ \lambda 0$ at the lowest resonant frequency.
- The ground plane length is optimized to enhance performance across the bands.

To address these challenges, this work integrates modified circular ring structures, optimized slot placements, and multiple CSSRR configurations. The proposed antenna demonstrates superior multi-band performance, improved impedance matching, and enhanced radiation characteristics. Future enhancements could involve further optimization of the CSSRR configurations and the integration of meta- surfaces to achieve broader bandwidths and increased efficiency in practical applications.

2. ANTENNA DESIGN EVOLUTION

Fig. 1 illustrates the step-by-step evolution of the proposed design, labeled as Antenna 1, Antenna 2, Antenna 3, and Antenna 4. Table I provides the detailed measurements and parameter values for each stage. Initially, the ring-shaped antenna design has an outer radius R_1=9.3 mm, an inner radius R_2=6.8 mm, and a width w_1=2.5 mm. The ring-shaped antenna integrated a rectangular patch at its bottom, measuring y=1.95 mm in length and x=0.75 mm in width. The overall antenna includes a full ground plane measuring 20×20 mm^2.

This configuration achieved a resonant band at 3.16 GHz, as calculated using Equation 1. The simulated results were validated against the theoretical circular patch Equation 1, and the design is labeled as Antenna 1 in Fig. 1.

Further, the rectangular slot has been etched from Antenna 1 to develop a multi-band antenna. Fig. 1 labels the slot as Antenna 2, with a width (Rwx) of 0.5 mm. The antenna forms this slot in the traditional CSSRR shape. The inclusion of this slot resulted in the formation of a penta-band with resonant frequencies at 1.54 GHz, 3.18 GHz, 4.46 GHz, 6.24 GHz, and 7.46 GHz. Among these, only the resonant frequency at 6.24 GHz exhibits excellent impedance matching, while the reflection coefficients of the remaining four frequencies indicate poor impedance matching.

Table 1. Antenna Dimensions with detail Parameters

Parameter	R_1	R_2	R_3	R_w	R_{wx}
Dimension (mm)	9.3	6.8	5.8	1.85	0.5
Parameter	W_1	X	Y	W_2	G
Dimension (mm)	2.5	0.75	1.95	1.5	0.7
Parameter	X_1	R_4	R_5	-	-
Dimension (mm)	1.5	3.6	0.5	-	-

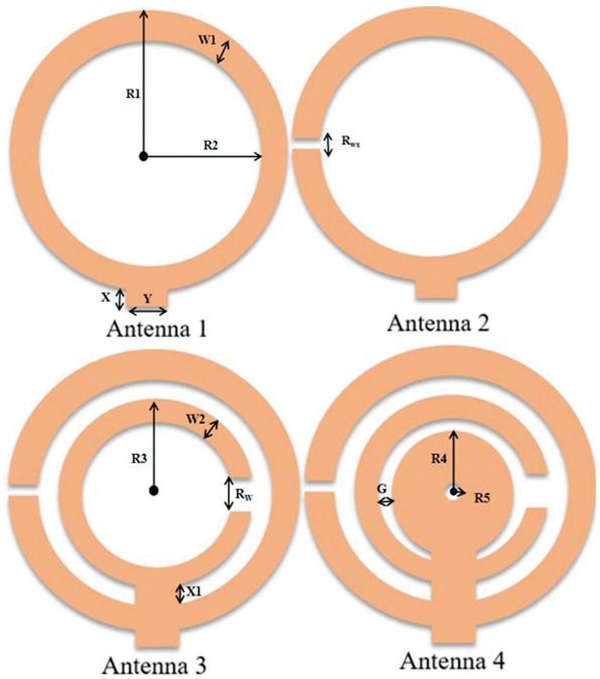


Fig. 1. The Evolution of the Proposed Antenna 1,2,3 and 4

In Antenna 3, a circular-shaped split ring resonator (CSSRR) with a radius (R_3) of 5.8 mm and a ring width (W_2) of 1.5 mm has been incorporated from Antenna 2. A rectangular slot with a length (y) of 1.85 mm and a width (x_1) of 1.5 mm has been connected to the first CSSRR. Additionally, a 1.85 mm-wide slot is etched from the second ring-shaped antenna to form the second CSSRR. This configuration resulted in a hexa-band antenna with resonant frequencies at 1.26 GHz, 2.22 GHz, 3.24 GHz, 4.06 GHz, 5.66 GHz, and 6.84 GHz. However, the lowest and highest frequencies exhibit poor impedance matching.

The circular-shaped antenna has a radius (R_4) of 3.6 mm, with a gap (G) of 0.7 mm for the second CSSRR. Additionally, a smaller radius (R_5) of 0.5 mm has been etched from the center of the main radiating element of R_4. This structure represents an optimized multi-band antenna with an improved performance. Through the integration of modified circular rings, CSRRs, and slots, the antenna achieves octa-band operation. The lowest frequency now exhibits better impedance matching compared to Antenna 3. The optimized antenna is shown in Fig. 1, labeled as Antenna 4.

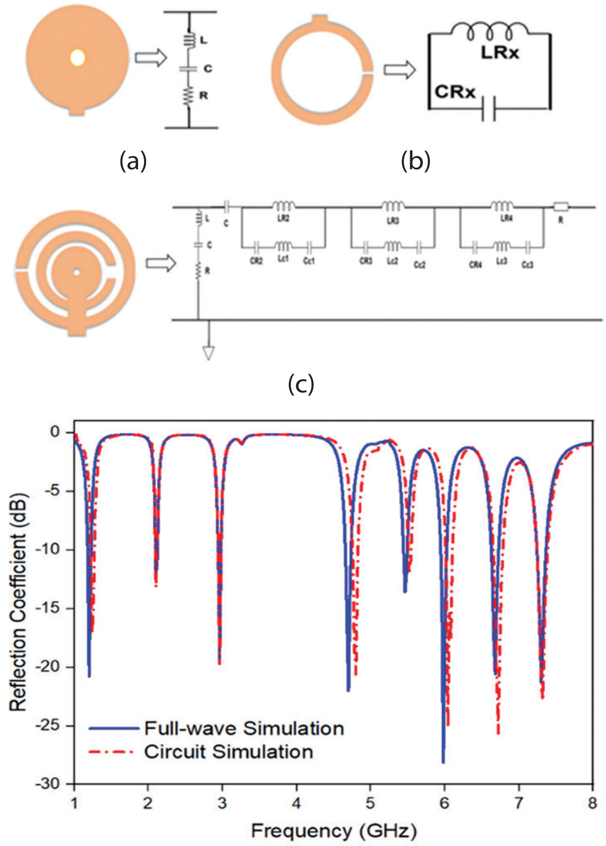
The radiator of the antenna is composed of two circular-shaped split ring resonator elements of structure with different dimensions on a single-layer square-shaped FR4 substrate Generally, the approximate original value for the radius of the circular-shaped split ring with a radius of R1; thus, the exact value of R1 can be evaluated with the assistance of the consecutive circle patch radius Equitation 1.

$$a = \frac{F}{\left\{1 + \frac{2h}{\pi \in_r F}\left[ln\left(\frac{\pi F}{2h} + 1.7726\right)\right]\right\}^{0.5}} \qquad (1)$$

where $F = \frac{8.791 \times 10^9}{f_r \sqrt{\in_r}}$

Where 'a' is equal to the circular patch radius R_1. The following section presents the parametric study of the proposed antenna resonated at frequencies of 1.2, 2.1, 2.9, 4.7, 5.4, 5.9, 6.6, and 7.3 GHz.

Equivalent circuit is designed to further understand the evolution of the antenna. The circular shaped antenna shown in Fig. 2(a) can be represented as a series LC circuit. However, due to the center hole and the conductive nature of the circular patch, an additional resistance is added and the antenna is modeled using RLC circuit [34]. When the radius of the circular hole is increased, the antenna forms a split-ring structure and the equivalent parallel LC circuit defines the antenna (this is most commonly seen in the existing literature). The novelty of the paper is the multi-band response using a compact antenna design and when multiple split-rings are introduced in a concentric manner, the equivalent circuit is a series combination of LC circuits along with the initial RLC circuit as shown in Fig. 2(c). The capacitors (CR1, CR2, CR3) are the shunt capacitors, while CC1, CC2 and CC3 are the coupling capacitors that exist due to the coupling between the adjacent rings. Between the CR capacitor and the CC capacitor, there exists a coupling inductance LC. Fig. 2(d) shows the frequency response of the equivalent circuit model of the proposed antenna. Due to the coupling capacitors and LC resonant circuits an octa band response is achieved. The circuit simulation results match well with the full-wave simulation results with < 5% deviation. The resonant frequencies achieved from the full-wave simulator are 1.2 GHz, 2.1 GHz, 2.9 GHz, 4.7 GHz, 5.4 GHz, 5.9 GHz, 6.6 GHz, and 7.3 GHz. The resonant frequencies obtained from the circuit- simulation are 1.24 GHz, 2.11 GHz, 2.9 GHz, 4.82 GHz, 5.48 GHz, 6.0 GHz, 6.72 GHz, and 7.32 GHz. Minor discrepancies can be attributed to the ideal circuit elements and material losses that are not considered in the circuit simulation.

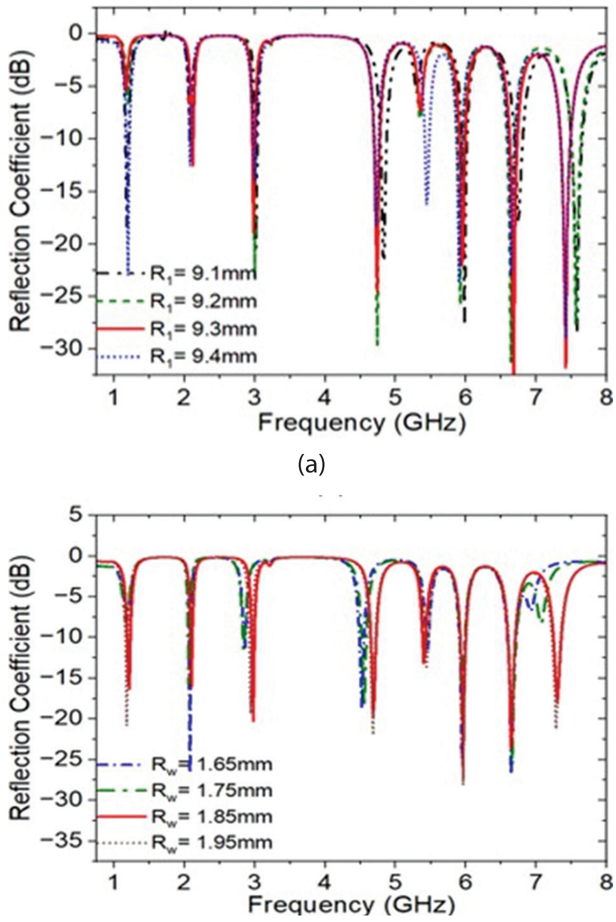


(d)

Fig. 2. The equivalent circuit of: (**a**) Circular-shaped antenna (**b**) CSSRR-shaped antenna, (**c**) Proposed model and (**d**) Comparison of the reflection coefficient of the proposed antenna with equivalent circuit response and full-wave simulation

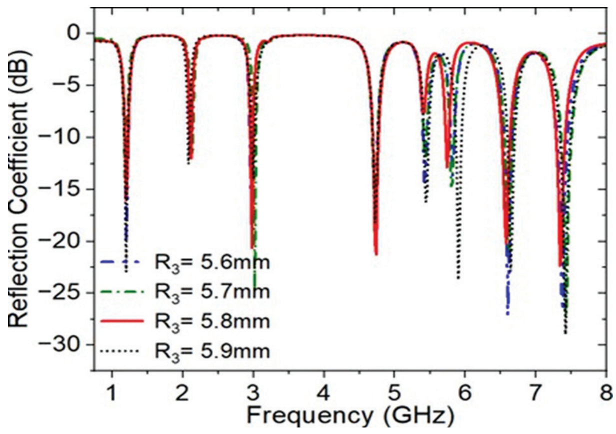
3. PARAMETRIC ANALYSIS AND WORKING MECHANISM

The octa-band antenna's structure and each band's principle have been analyzed. The antenna has been simulated using commercial EM solver to examine the impact of various settings on matching, frequency shift, bandwidth, and gain. Fig. 3 and Fig. 4 depicts the impact of parameter adjustments on the reflection coefficient. Fig. 3(a) displays S_{11} with the effect of varying R_1 from 9.1 mm to 9.4 mm. The desired frequency with good return loss is attained only when $R_1 = 9.3$ mm . In the other cases, i.e., R_1 =9.1 mm, 9.2 mm, and 9.4 mm, the Octa-band response is not detected. The parametric analysis of the R_w parameter, as shown in Fig. 3(b), illustrates a variation from 1.65 mm to 1.95 mm. Based on the observation of the S11 results, the optimized R_w value is determined to be 1.85 mm. In the other cases i.e. at 1.65 mm, 1.75 mm, and 1.95 mm a frequency shift and impedance mismatch are observed. Fig. 4(a) shows the simulated S_{11} results by increasing the radius R_3 from 5.6 mm to 5.9 mm. The necessary band resonance occurs only when optimized R_3=5.8 mm, while other values result in a frequency shift in the higher bands. The Fig. 4(b) with S_{11} represents the influence of R_{wx} from 0.4 mm to 0.7 mm on the major radiating patch. Due to this, there is a minor difference in the upper three bands. This analysis tries to determine the suitable value of R_{wx} to get the accurate octa band.

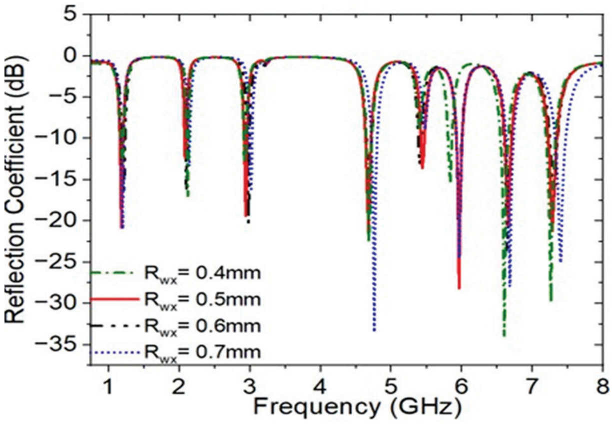


(b)

Fig. 3. Parametric analysis showing (**a**) Variation in R_1 (**b**) Variation in R_w



(a)



(b)

Fig. 4. Parametric analysis showing (**a**) Variation in R_3, (**b**) Variation in R_{wx}

Fig. 5 illustrates the surface current distribution at eight operational frequencies, revealing distinct patterns across the various frequencies.

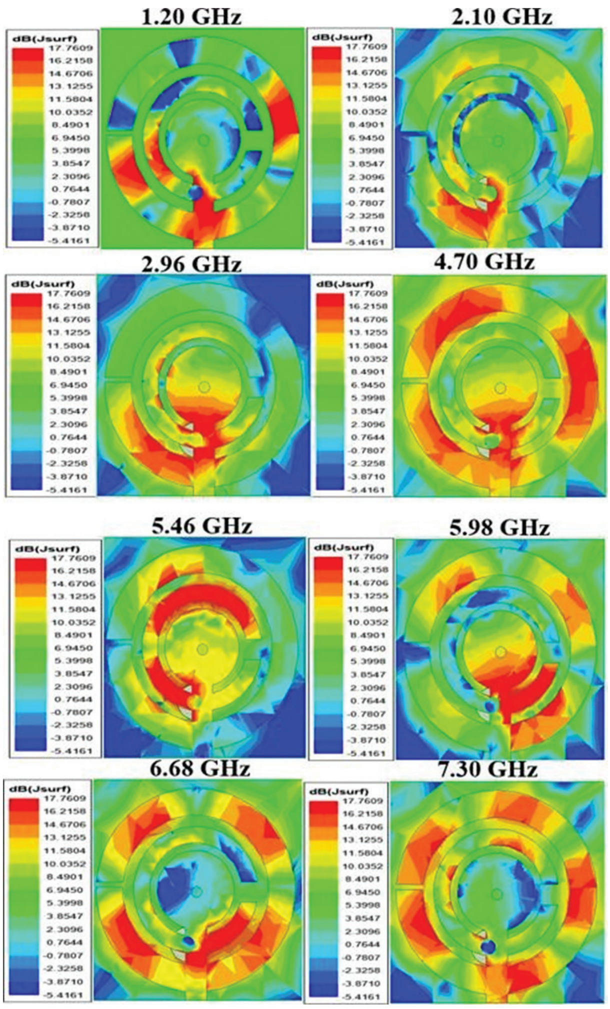


Fig.5. Surface current distribution of the resonating frequencies of the CSSRR antenna

At 1.20 GHz, the current is strong in the outer split-ring radiating area but weak in the central radiating area. At 2.10 GHz, the current is concentrated in the first and second split rings, with a small amount extending to the center. At 2.96 GHz, the current is intense around the feeding area and the right- side radiating area, while it is moderate on the opposite side. At 4.70 GHz and 6.68 GHz, the surface current is moderate in the central area and strong in the outer split- ring resonator. At 5.44 GHz, the current is strong in the inward split-ring resonator and weak in the outward split- ring resonator. At 5.84 GHz, the current is highly concentrated around the feeding position on the right side. Finally, at 6.68 GHz and 7.30 GHz, the surface current distribution is strong in the outer ring. These observations provide valuable insights into the antenna's behavior at different operational frequencies.

4. RESULTS AND DISCUSSION

To assess the effectiveness of the designed antenna, a prototype was fabricated and Fig. 6 presents the optimized antenna design with dimensions, including both the front and rear views, alongside the fabricated model.

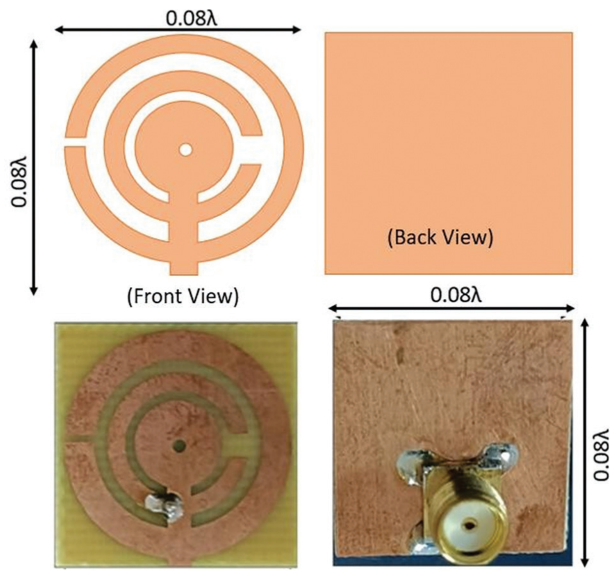


Fig. 6. Front view and Back view of the fabricated prototype

The calibration and measurement process of the reflection coefficient of the proposed CSSRR antenna using the E5063A model of a vector network analyzer (VNA), operating within the frequency range of 100 kHz to 15 GHz is shown in Fig. 7.

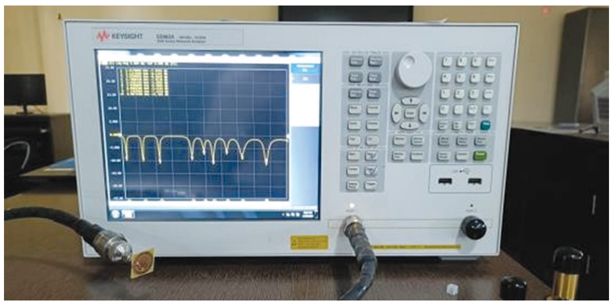


Fig. 7. Measuring the S-parameters using E5063A model of VNA and ranging 100KHz to 15 GHz

The reflection coefficient of the antenna has been verified with the simulated and measured results compared in Fig. 8. The operating bands are identified as 1.20, 2.10, 2.96, 4.70, 5.46, 5.98, 6.68, and 7.30 GHz. The reflection coefficient is below - 10 dB, indicating good performance, with only 10% of the power reflected and 90% radiated by the antenna.

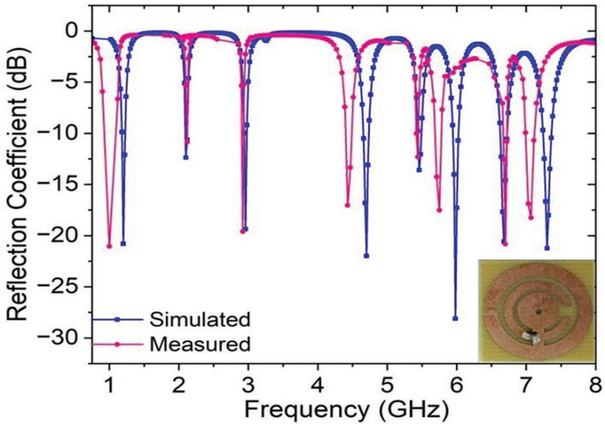


Fig. 8. Simulated and measured results of the proposed antenna design

The impedance bandwidth for each resonance of the antenna is shown in Table 2.

Table 2. Impedance Matching and Return loss Results

Resonance Frequency (*fr*) in GHz	Impedance Bandwidth (*IBW*) in GHz	Reflection Coefficient (S_{11}) in dB
1.22	1.1-1.25	-22
2.12	2.1-2.2	-13
3	2.9-3.2	-21
4.76	4.55-4.8	-23
5.4	5.3-5.45	-14
5.94	5.8-6.25	-28
6.70	6.6-6.78	-23
7.74	7.3-7.6	-24

The simulated and measured VSWR (Voltage Standing Wave Ratio) values are shown in Fig. 9. They are both below 2 and within the acceptable range for real-time applications in the relevant operating bands. At certain resonant frequencies, the VSWR is observed to be less than 1.5. The simulated and measured results were are in close agreement with each other.

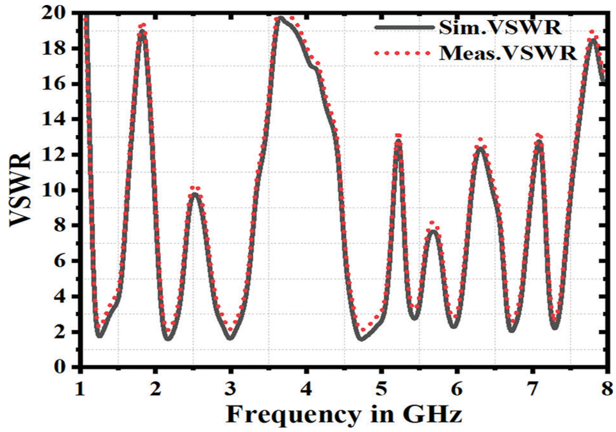
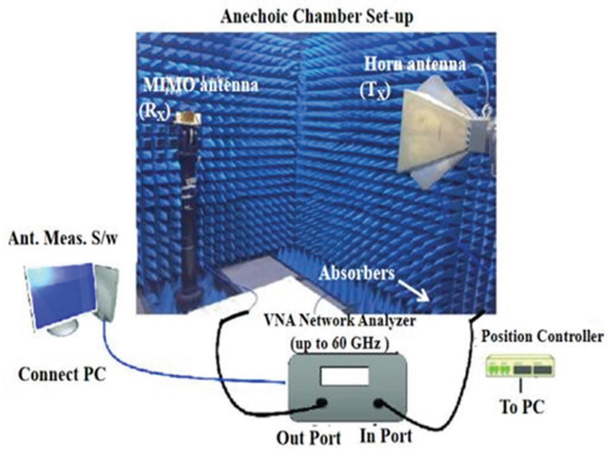


Fig. 9. Simulated and measured VSWR

Among the operational frequencies, the far-field patterns are measured for the frequencies where maximum impedance matching is achieved. Fig. 10 (a) illustrates the anechoic chamber setup from the inside, while Fig. 10 (b) shows the exterior setup used for measuring the radiation pattern and gain of the proposed antenna.



(a)

(b)

Fig. 10. Radiation and gain measuring setup in anechoic Chamber: (a) inside Setup (b) outside setup

The 2D radiation patterns of the E-plane ($\phi = 0^0$) and H-plane (($\phi = 90^0$) corresponding to the resonating frequencies at (a) 1.22 GHz, (b) 2.12 GHz, (c) 3.0 GHz, (d) 4.76 GHz, (e) 5.40 GHz, (f) 5.94 GHz, (g) 6.70 GHz, and

(h) 7.44 GHz, as shown in Fig. 11. The figure illustrates both the simulated and measured radiation patterns for these frequencies. The patterns demonstrate stable behavior, ensuring minimal cross-polarization and focused radiation in specific directions.

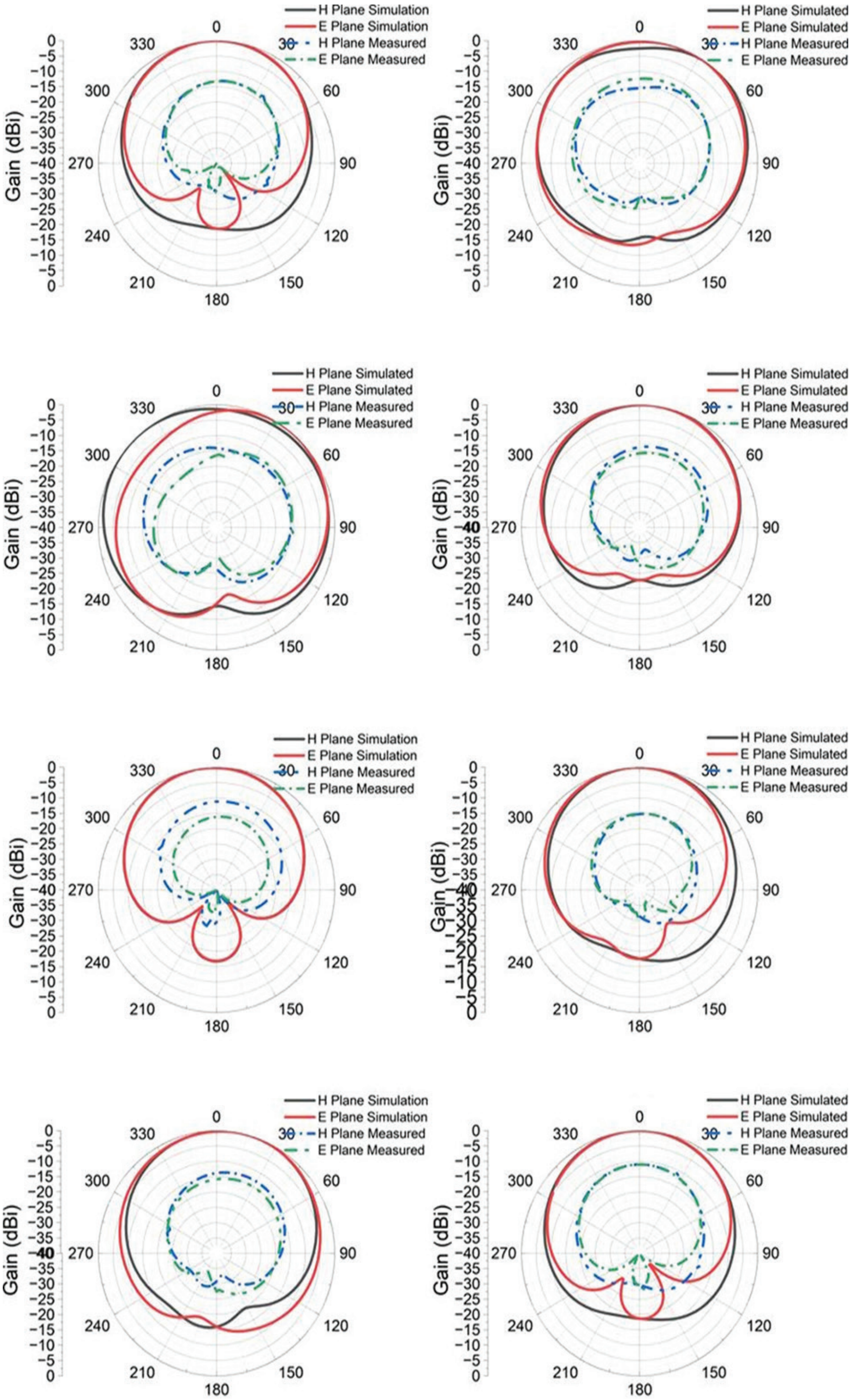


Fig. 11. Radiation pattern for (a) 1.2 GHz (b) 2.1 GHz (c) 2.9 GHz (d) 4.7 GHz (e) 5.4 GHz (f) 5.9 GHz (g) 6.6 GHz and (h) 7.3 GHz

Fig. 12. presents the simulated and measured peak gain and radiation efficiency of the proposed antenna design. The results demonstrate improved performance, with variations observed across different operating bands. The radiation efficiency ranges from 77% to 85% for the octa-band configuration of the proposed antenna. Table 3 provides the far-field simulated and measured results.

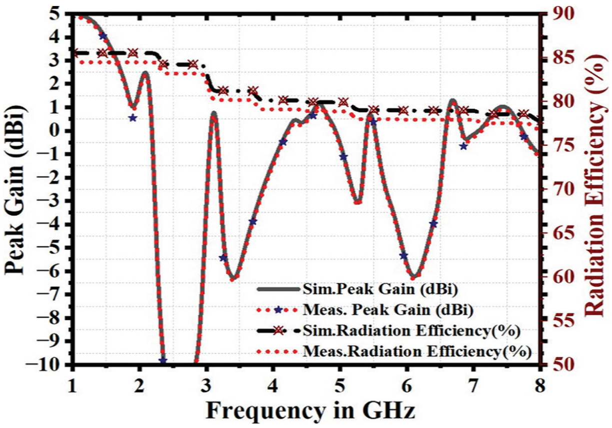


Fig. 12. Simulated and measured Peak Gain along with the Radiation Efficiency results of proposed design

Table 3. Simulated and measured peak gain and radiation efficiency at the corresponding resonant frequencies

f_r in GHz	Gain (dBi) (Sim.)	Gain (dBi) (meas)	Eff (Sim)	Eff (Meas)
1.22	4.80	4.71	85.10	84.27
2.12	2.81	2.68	85.44	84.46
3.0	0.98	0.86	81.2	80.17
4.76	1.39	1.16	79.9	78.87
5.40	1.167	1.041	79.2	77.99
5.94	0.50	0.377	79	77.99
6.70	0.30	0.17	78.5	77.92
7.44	1.04	0.92	78.5	77.52

A comprehensive analysis of the proposed antenna design compared to state-of-the-art models demonstrates its significant advancements in key performance areas. The proposed design excels in achieving enhanced multiband operation, better impedance matching, and improved integration potential with modern wireless systems, all while maintaining a compact and low-profile structure out of all the existing antennas. The findings of this research are summarized in Table 4 and they illustrate the innovative contributions of the design and it can be observed that the challenges such as achieving simultaneous miniaturization, multi- band operation and stable gain are resolved by this proposed antenna. This study highlights the effectiveness and practicality of the antenna, marking it as a valuable solution for advanced wireless communication applications.

5. CONCLUSION

The proposed antenna design effectively achieves octa- band frequency response through a compact radiator based on a circular-shaped split ring resonator (CSSRR) structure. Operating at 1.22, 2.12, 3, 4.76, 5.4, 5.94, 6.70, and 7.44 GHz, the antenna exhibits excellent radiation characteristics, with a peak gain of 4.71 dBi and radiation efficiency ranging from 74% to 85% across all bands. Its compact size, cost-effectiveness, and superior performance make it an ideal solution for sub-6 GHz 5G applications and other wireless communication systems, including Wi-Fi, WLAN, and WiMAX. Furthermore, it has the potential to support higher frequency bands, while maintaining optimized efficiency and gain. This makes the antenna suitable for emerging use cases in autonomous vehicles, smart cities, and defense systems. Overall, the proposed antenna provides a strong platform to address the evolving needs of modern and next-generation wireless communication systems.

Table 4. Comparative study between the state-of-the-art work of the proposed design

Ref.	Area ($\lambda 0 \times \lambda 0$)	No. of Bands	Freq. Ratio	Freq. band (fH – fL)	Gain (dBi)	Eff(%)
7	1.16 × 1.16	6	2.01	11.2	1.34	90
9	0.28 × 0.24	6	2.84	3.87	2.81	92
10	0.29 × 0.29	6	2.05	2.63	-	-
11	0.35 × 0.14	6	2.53	3.63	7.39	-
12	0.55 × 0.55	6	5.58	23.46	2.3	71
13	0.57 × 0.57	6	7.51	22.55	3.23	87
14	0.27 × 0.27	7	4.25	5.6	-	-
15	1.17 × 1.17	7	2.91	53.9	7.16	99
29	0.26 × 0.178	8	7.25	4.5	4	-
TW	0.08 × 0.08	8	6.08	6.1	4.71	85

6. REFERNCES

[1] R. Saleem, M. Bilal, H. T. Chattha, A. Mushtaq, M. F. Shafique, "An FSS based multiband MIMO system incorporating 3D antennas for WLAN/ WiMAX/5G cellular and 5G Wi-Fi applications", IEEE Access, Vol. 7, 2019, pp. 144732-144740.

[2] Y. Liu, L. Shafai, D. Isleifson, C. Shafai, "Split Ring Antennas and Their Application for Antenna Miniaturization", Sensors, Vol. 23, No. 2, 2023, p. 846.

[3] R. A. Mellita, S. S. Karthikeyan, P. Damodharan, D. S. Chandu, "A miniaturized quad-band frequency selective surface for C-band applications", Journal of Electromagnetic Waves and Applications, Vol. 35, No. 14, 2021.

[4] R. K. Saraswat, M. Kumar, "Implementation of an Octa-band Metamaterial Inspired Frequency Band Reconfigurable Antenna for IoT applications in Wireless Standards", IETE Journal of Research, Vol. 70, No. 5, 2022.

[5] S. E. Kilani, L. E. Abdellaoui, J. Zbitou, M. Latrach, "A compact dual band PIFA antenna for GPS and ISM BAND applications", Indonesian Journal of Electrical Engineering and Computer Science, Vol. 14, No. 3, 2019, pp. 1266-1271.

[6] M. Thirupathi, B. Harikrishna, "Reduced mutual coupling multiband MIMO patch antenna with swastik type mushroom EBG", Indonesian Journal of Electrical Engineering and Computer Science, Vol. 14, No. 1, 2019, pp. 490-494.

[7] A. Sabah, M. J. Frhan, "A new patch antenna for ultra wide band communication applications", Indonesian Journal of Electrical Engineering and Computer Science, Vol. 18, No. 2, 2020, pp. 848-855.

[8] S. Liu, Z. Wang, Y. Dong, "Compact wideband SRR-inspired antennas for 5G microcell applications", IEEE Transactions on Antennas and Propagation, Vol. 69, No. 9, 2021, pp. 5998-6003.

[9] V. Rajeshkumar, R. Rajkumar, "SRR loaded compact tri-band MIMO antenna for WLAN/WiMAX applications", Progress in Electromagnetics Research Letters, Vol. 95, 2021, pp. 43-53.

[10] A. El Yousfi, A. Lamkaddem, K. A. Abdalmalak, D. Segovia-Vargas, "A miniaturized triple-band and dual-polarized monopole antenna based on a CSRR perturbed ground plane", IEEE Access, Vol. 9, 2021, pp. 164292-164299.

[11] S. P. J. Christydass, N. Gunavathi, "Dual-band complementary split-ring resonator engraved rectangular monopole for GSM and WLAN/ WiMAX /5G sub-6 GHz band (new radio band)", Progress In Electromagnetics Research C, Vol. 113, 2021, pp. 251-263.

[12] S. Ullah, S. Ahmad, J. A. Flint, "A multi-band switchable antenna for Wi-Fi, 3G advanced, WiMAX, and WLAN wireless applications", International Journal of Microwave and Wireless Technologies, Vol. 10, No. 8, 2018, pp. 991-997.

[13] P. Chaurasia, B. Kanaujia, S. Dwari, M. K. Khandelwal, "Antenna with hexa-band capabilities for multiple wireless applications", Progress In Electromagnetics Research C, Vol. 82, 2018, pp. 109-122.

[14] V. Singh, B. Mishra, A. K. Dwivedi, R. Singh, "Inverted l-notch loaded hexa band circular patch antenna for x, ku/k band applications", Microwave and Optical Technology Letters, Vol. 60, No. 8, 2018, pp. 2081-2088.

[15] J. O. Abolade, D. B. Konditi, V. M. Dharmadhikary, "Compact hexa-band bio-inspired antenna using asymmetric microstrip feeding technique for wireless applications", Heliyon, Vol. 7, No. 2, 2021.

[16] R. Dhara, T. Kundu, M. Sanoj, "Circularly polarized hexa-band planar slot antenna for WLAN and wireless communication application", Radioelectronics and Communications Systems, Vol. 65, No. 1, 2022, pp. 48-59.

[17] D. Tiwari, J. A. Ansari, A. K. Saroj, M. Kumar, "Analysis of a miniaturized hexagonal sierpinski gasket fractal microstrip antenna for modern wireless communications", AEU-International Journal of Electronics and Communications, Vol. 123, 2020, p. 153288.

[18] R. Sachan, D. C. Dhubkarya, "Photonic bandgap hepta-band stacked microstrip antenna for l, S and C band applications", Wireless Personal Communications, Vol. 116, No. 3, 2021, pp. 1913-1931.

[19] W. Shahjehan, I. Hussain, I. Ali, A. Riaz, P. Uthansakul, "Hepta-band antenna for 5G applications", Wireless Personal Communications, Vol. 125, No. 3, 2022, pp. 2031-2054.

[20] D. Huang, Y. Wang, "Eight-band antenna for full-screen metal frame LTE mobile phones", IEEE Trans. Antennas Propagation, Vol. 67, No. 3, 2018, pp. 1527-1534.

[21] Y. Wang, Z. Du, "Wideband monopole antenna with less nonground portion for octa-band WWAN/LTE mobile phones", IEEE Transactions on Antennas and Propagation, Vol. 64, No. 1, 2015, pp. 383-388.

[22] D. Huang, Y. Wang, "An octa-band monopole antenna with a small nonground portion height for LTE/WWAN mobile phones", IEEE Transactions on Antennas and Propagation, Vol. 65, No. 2, 2016, pp. 878-882.

[23] K.-L. Wong, P.-R. Wu, "Dual-wideband linear open slot antenna with two open ends for the LTE/WWAN smartphone", Microwave and Optical Technology Letters, Vol. 57, No. 6, 2015, pp. 1269-1274.

[24] D. Huang, Z. Du, "Eight-band antenna with a small ground clearance for LTE metal-frame mobile phone applications", IEEE Antennas and Wireless Propagation Letters, Vol. 17, No. 1, 2017, pp. 34-37.

[25] R. Tang, Z. Du, "Wideband monopole without lumped elements octa-band narrow-frame LTE smartphone", IEEE Antennas and Wireless Propagation Letters, Vol. 16, 2016, pp. 720-723.

[26] D. Huang, Z. Du, Y. Wang, "Compact octa-band monopole antenna with independently tuning for WWAN/LTE mobile phones", Microwave and Optical Technology Letters, Vol. 59, No. 1, 2017, pp. 208-214.

[27] K.-L. Wong, C.-Y. Tsai, "Half-loop frame antenna for the LTE metal casing tablet device", IEEE Transactions on Antennas and Propagation, Vol. 65, No. 1, 2016, pp. 71-81.

[28] Z.-Q. Xu, Q.-Q. Zhou, Y.-L. Ban, S. S. Ang, "Hepta-band coupled fed loop antenna for LTE/WWAN unbroken metal-rimmed smartphone applications", IEEE Antennas and Wireless Propagation Letters, Vol. 17, No. 2, 2017, pp. 311-314.

[29] Y. Liu, W. Cui, Y. Jia, A. Ren, "Hepta-band metal-frame antenna for LTE/WWAN full-screen smartphone", IEEE Antennas and Wireless Propagation Letters, Vol. 19, No. 7, 2020, pp. 1241-1245.

[30] D. Wu, S. Cheung, T. Yuk, "A compact and low- profile loop antenna with multiband operation for ultra-thin smart phones", IEEE Transactions on Antennas and Propagation, Vol. 63, No. 6, 2015, pp. 2745-2750.

[31] Ab Rashid, A. Hafizun, B. H. Ahmad, M. Z. A. Abd Aziz, "CPW Fractal Antenna with Third Iteration of Pentagonal Sierpinski Gasket Island for 3.5 GHz WiMAX and 5.2 GHz WLAN Applications", International Journal of Electrical and Computer Engineering Systems, Vol. 14, No. 2, 2023, pp. 129-134.

[32] M. Abbasi, T. A. Latef, "Wideband circularly polarized fractal antenna with SSRR metasurface for 5G applications", International Journal of Electrical and Computer Engineering Systems, Vol. 15, No.1, 2024, pp. 89-98.

[33] K. Vinayagam, R. Natarajan, "Polarization Reconfigurable Patch Antenna Using Parasitic Elements for Sub-6 GHz Applications", International Journal of Electrical and Computer Engineering Systems, Vol. 16, No. 1, 2025, pp. 1-7.

[34] F. Liu, J. Guo, Z. Zhao, L. Zhao, G.-L. Huang, Y. Li, Y. Yin, "Coupled meander-line resonators for mutual coupling suppression of microstrip patch antennas", International Journal of RF and Microwave Computer-Aided Engineering, Vol. 31, No. 3, 2021, p. e22528.

Bibliometric Analysis of Scientific Production of Intelligent Video Surveillance

Review Paper

Wagner Vicente Ramos*
Universidad Nacional de Cañete
Faculty of Engineering, School of Systems Engineering
San Agustín Street No. 124, San Vicente de Cañete, Lima, Perú
wvicente@undc.edu.pe

Alex Pacheco Pumaleque
Universidad Nacional de Cañete
Faculty of Engineering, School of Systems Engineering
San Agustin Street No. 124, San Vicente de Cañete, Lima, Perú
apacheco@undc.edu.pe

Jhonny Gavino Torres
Universidad Nacional de Cañete
Graduate School, Master's Degree in Higher Education and Research
Av. Mariscal Benavides 1370 House of Culture, San Vicente de Cañete Lima, Peru
44326021@undc.edu.pe

*Corresponding author

Abstract *– This article offers a bibliometric analysis of academic research in intelligent video surveillance, evaluating its evolution between 2000 and 2024. 1,343 documents were collected from the Scopus database and the PRISMA methodology was applied to organize the search and selection of relevant publications. The findings show a notable increase in the number of studies, reaching its highest point in 2022, driven by advances in artificial intelligence, the Internet of Things (IoT) and deep learning. China leads scientific production in this field, followed by India and the United States. Main research areas include real-time surveillance using deep learning methods, sequential and transfer learning techniques, as well as the use of advanced YOLO, Faster-RCNN and RFCN algorithms in controlled environments; however, detecting unusual behavior is a latent challenge.*

Keywords: *video, iot, cybersecurity, surveillance, behavioral detection*

Received: November 1, 2024; Received in revised form: January 24, 2025; Accepted: January 25, 2025

1. INTRODUCTION

Smart video surveillance has emerged as a key tool in the protection and monitoring of public spaces, not only because of its ability to prevent crime, but also because of its suitability to process large volumes of information in real time. These technologies provide proactive detection of suspicious or dangerous behavior, facilitating quick and effective responses from authorities. A central aspect of intelligent video surveillance is its integration with deep learning algorithms and cloud-based systems or fog computing, which allow real-time image processing without overloading traditional cloud infrastructures [1]. This capability significantly improves safety, as intelligent systems can identify objects or people in potentially dangerous situations, improving incident response capability [2]. In addition, the use of artificial intelligence in these systems not only optimizes the efficiency of surveillance, but also presents the possibility of respecting privacy through algorithms designed to anonymize individuals in situations where no immediate risks are detected. However, it is crucial to take into account the public perception regarding privacy and the associated risks, as studies have shown that a considerable percentage of the population perceives these systems as invasive, which can affect their acceptance in society [2].

Smart video surveillance is an advanced system leveraging artificial intelligence technologies, such as deep learning, to enhance event detection and enable real-time decision-making. By integrating sensors and cameras with Internet of Things (IoT) networks, these systems can automatically identify suspicious behavior or anomalies within a given environment. This approach is crucial for improving security in urban areas, enabling authorities to respond swiftly to potential threats [1]. In addition, implementing fog computing in these systems reduces latency and improves data processing efficiency, which is essential for real-time surveillance applications [3].

A key aspect of smart video surveillance is its capacity for continuous learning and improvement. Machine learning algorithms in these systems can be regularly trained on new data, enhancing their accuracy in detecting incidents, even in complex scenarios or environments with multiple elements. These capabilities not only make it possible to identify objects or people, but also to predict unusual behaviors before they become a serious problem [2]. In addition, advanced algorithms based on the behavior of tuna swarms, combined with deep learning techniques, have proven effective in identifying violent acts in real-time within surveillance platforms [4].

In addition to security aspects, smart surveillance systems must face ethical challenges, especially in terms of respect for privacy and the management of personal data. Public acceptance of these systems depends to a large extent on how the balance between security and the protection of individual privacy is managed. Concerns about possible misuse of the collected data are common, forcing developers to incorporate various actions that ensure the processing of data that is actually necessary for security purposes (Golda et al., 2022). For this, it is essential to establish a legal framework that clearly regulates the collection, processing, and protection of personal information [5].

Smart video surveillance has been implemented internationally in various contexts to improve public safety and optimize the response to critical situations. An example of its application is the use of surveillance systems based on fog computing in Saudi Arabia, where an intelligent monitoring system has been deployed for the recognition of suspicious actions in real time. This system uses distributed cameras and advanced image processors to detect potential threats and generate alerts to the competent authorities, thus achieving a faster and more effective response [1]. Another significant example is in Germany, where public perceptions of smart video surveillance systems in public spaces were assessed. These systems have been used to detect risks in mass events, such as concerts or protests, where automated monitoring makes it possible to identify dangerous behaviors and prevent safety incidents before they occur [2]. These cases reflect how smart video surveillance has been adapted to local needs, improving both public safety and social acceptance in different countries.

This analysis of publications indexed in databases such as Scopus aims to identify the primary algorithms based on deep neural networks that enable the development of real-time intelligent video surveillance systems, particularly for object detection and tracking. Additionally, it seeks to evaluate the impact and evolution of research in the field of Intelligent Video Surveillance.

2. METHOD

2.1. SEARCH STRATEGY

This study follows a bibliometric methodology, based on the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) approach. The data was obtained from the Scopus database on September 3, 2024, using the search terms: TITLE ("Video surveillance" AND "Smart"). Scopus was selected as the primary database for this study due to its extensive coverage of scientific publications in various disciplines, as well as its ability to provide comprehensive and consistent metadata. Instead, Web of Science was discarded, despite its wide access to high-quality academic research, as most of the journals it indexes (approximately 99%) are already included in Scopus, which would lead to redundant data collection without adding additional value to bibliometric analysis [6]. Google Scholar was also excluded due to its limited ability to provide the specific information needed for bibliometric analysis. Although Google Scholar is a useful tool for general literature searches, it lacks advanced features such as citation display and precise categorization by document type and relevance, making it inappropriate for studies that require a more rigorous and systematic bibliometric approach [7].

The flowchart (Fig. 1) shows the process of selecting the documents. The document selection process began with the identification of 1,406 records extracted from the Scopus database, based on specific search terms. Various exclusion criteria were then applied to purify the initial sample.

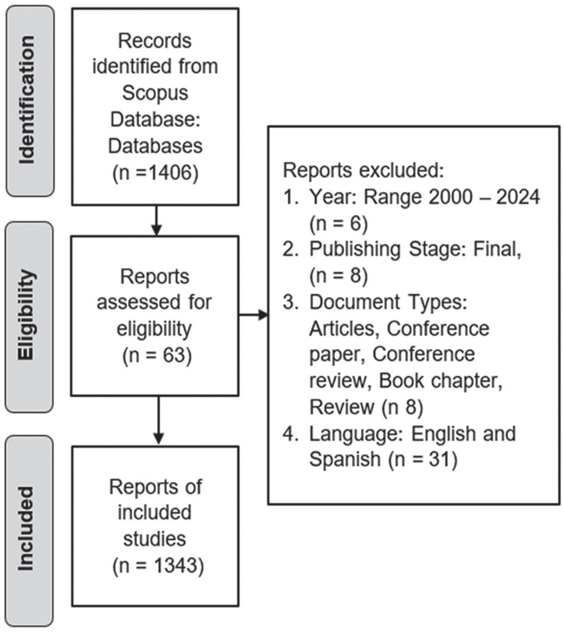


Fig. 1. Graphical representation of the record filtering process based on the PRISMA approach

First, six documents that did not correspond to the time range defined for the study, which spanned from the year 2000 to September 3, 2024, were discarded.

Eight studies that were not at a final publication stage, i.e. those that were under review or had not yet been fully published, were also removed.

Then, a more detailed selection was made according to the type of document, prioritizing scientific articles, conference proceedings, reviews and book chapters, while other non-relevant formats were excluded. Finally, in the selection by language, 31 publications written in languages other than Spanish or English were eliminated, since only these two languages were considered suitable for the present analysis. After this rigorous selection process, a total of 1,343 documents were retained for bibliometric analysis, complying with the previously established inclusion criteria.

2.2. DATA ANALYSIS

The collected data was organized and analyzed using the open-source software "bibliometrix R-package", widely recognized for its ability to process and visualize bibliometric data efficiently. This tool allowed an exhaustive analysis of scientific production in intelligent video surveillance, identifying the main research trends and mapping collaboration networks between authors, institutions and countries. Through the Biblioshiny graphical interface, time evolution graphs were generated that evidenced the significant growth of publications from 2017 onwards, driven by advances in artificial intelligence and the Internet of Things (IoT). Additionally, co-citation analysis and keyword co-occurrence techniques were used to identify the relationships between the most prolific researchers and emerging topics in this field. The results were presented in tables and figures, providing an overview of the current state of research, the predominant areas and the most relevant international collaborations, making it easier to identify development patterns and possible areas of future study.

3. RESULTS AND DISCUSSION

3.1. EVOLUTION OF SCIENTIFIC PRODUCTION

In Fig. 2, the production of 1,343 scientific articles generated during a period of 25 years (2000 to September 3, 2024) is shown.

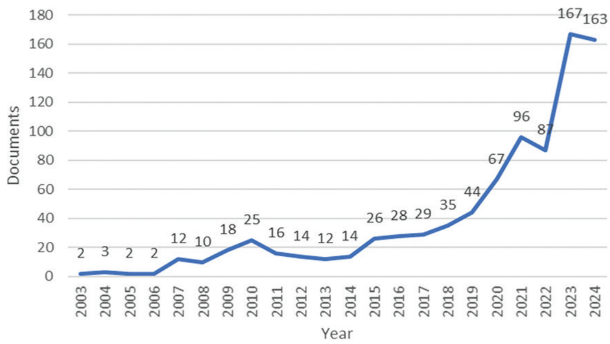


Fig. 2. Growth in the number of annual publications

The graph shows that between 2000 and 2016, scientific production in intelligent video surveillance remained moderate and stable, with slight fluctuations that did not exceed 52 annual publications. This suggests that, during this period, the technology was still in an emerging phase, facing technological limitations and lack of infrastructure for its mass implementation. As of 2017, there has been an exponential growth in the number of publications, reaching a peak in 2022 with 148 documents, reflecting the growing interest driven by advances in artificial intelligence and big data, which have been fundamental for the development of intelligent surveillance systems. However, in 2024 there is a significant decrease in the number of publications, with only 79 documents to date, which could be related to the saturation of the topic or the shift towards new technological areas such as differential privacy [8].

3.2. MORE PRODUCTIVE JOURNALS

Fig. 3 shows that 25.24% of the total publications were issued by the top 20 journals, which shows an uneven distribution.

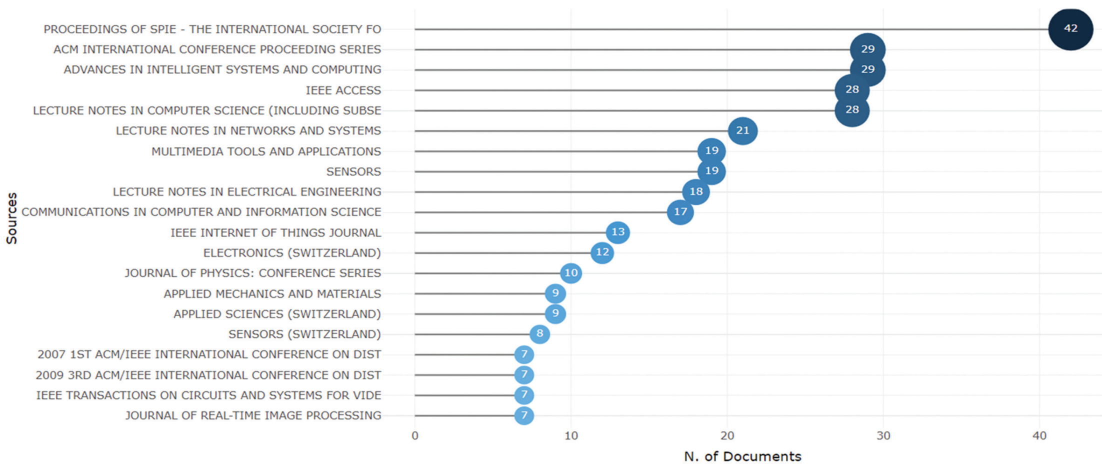


Fig. 3. Most relevant sources on smart video surveillance

Taking into account PROCEEDINGS OF SPIE - THE INTERNATIONAL SOCIETY FOR OPTICAL ENGINEERING (42) the source with the most documents available, followed by ACM INTERNATIONAL CONFERENCE PROCEEDING SERIES (29), ADVANCES IN INTELLIGENT SYSTEMS AND COMPUTING (29), IEEE ACCESS (28), LECTURE NOTES IN COMPUTER SCIENCE (INCLUDING SUBSERIES LECTURE NOTES IN ARTIFICIAL INTELLIGENCE AND LECTURE NOTES IN BIOINFORMATICS) (28), LECTURE NOTES IN NETWORKS AND SYSTEMS (21), MULTIMEDIA TOOLS AND APPLICATIONS (19), SENSORS (19), LECTURE NOTES IN ELECTRICAL ENGINEERING (18), COMMUNICATIONS IN COMPUTER AND INFORMATION SCIENCE (17), IEEE INTERNET OF THINGS JOURNAL (4), IEEE INTERNET OF THINGS JOURNAL (13), ELECTRONICS (SWITZERLAND) (12), JOURNAL OF PHYSICS: CONFERENCE SERIES (10), APPLIED MECHANICS AND MATERIALS (9), APPLIED SCIENCES (SWITZERLAND) (9), SENSORS (SWITZERLAND) (8), 2007 1ST ACM/IEEE INTERNATIONAL CONFERENCE ON DISTRIBUTED SMART CAMERAS, ICDSC (7), 2009 3RD ACM/IEEE INTERNATIONAL CONFERENCE ON DISTRIBUTED SMART CAMERAS, ICDSC 2009 (7), IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY (7) y JOURNAL OF REAL-TIME IMAGE PROCESSING (7). In general terms, the set of sources indicates that research in this field is highly favored by the integration of various disciplines, such as electrical engineering, computing, artificial intelligence and optics. International congresses and conferences, especially those organized by SPIE and ACM, are essential for the dissemination and exchange of knowledge, which justifies their relevance in the graphic. This underlines the key role played by these specialized forums in driving the progress and dissemination of technologies linked to Smart Video Surveillance.

3.3. MORE PRODUCTIVE JOURNALS

Fig. 4 shows the 20 most influential authors based on the number of articles in Scopus. The leading in the number of publications are NA NA, Wang Y, Chen Y, Zhang Y, Rinner B, Wang S, Xu Z, Zhang X, Ahamed I, Giorgi R, Li H, Li J, Li Y, Liu Y, Tian Y, Zhang J, Hampapur A, Singh S, Wang L, Wang X. The analysis of the writings of these renowned authors provides researchers with a comprehensive view that facilitates a deeper understanding of the field of study, as well as a critical reflection on the methodologies employed in their own research.

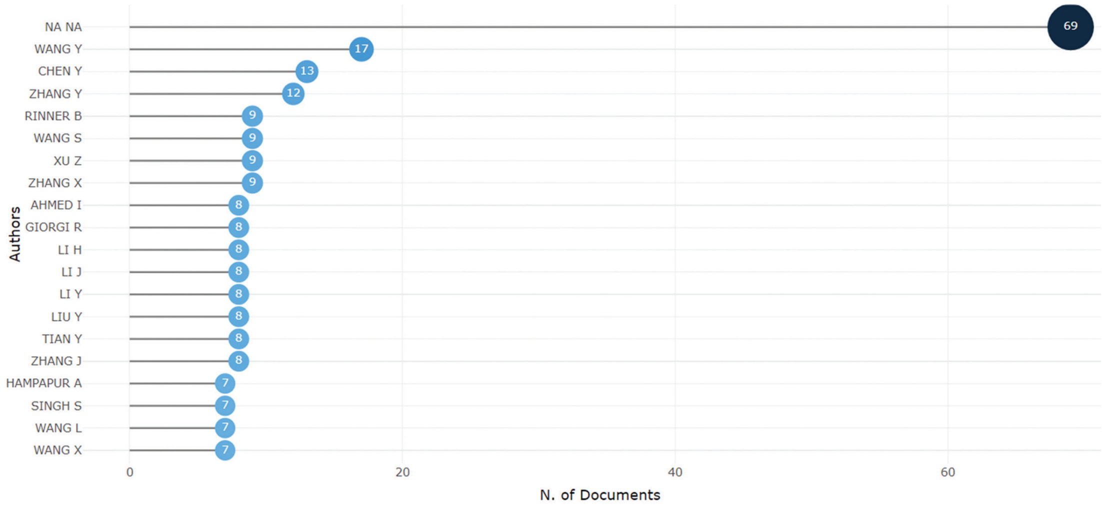


Fig. 4. Authors with the greatest influence according to the number of publications in Scopus

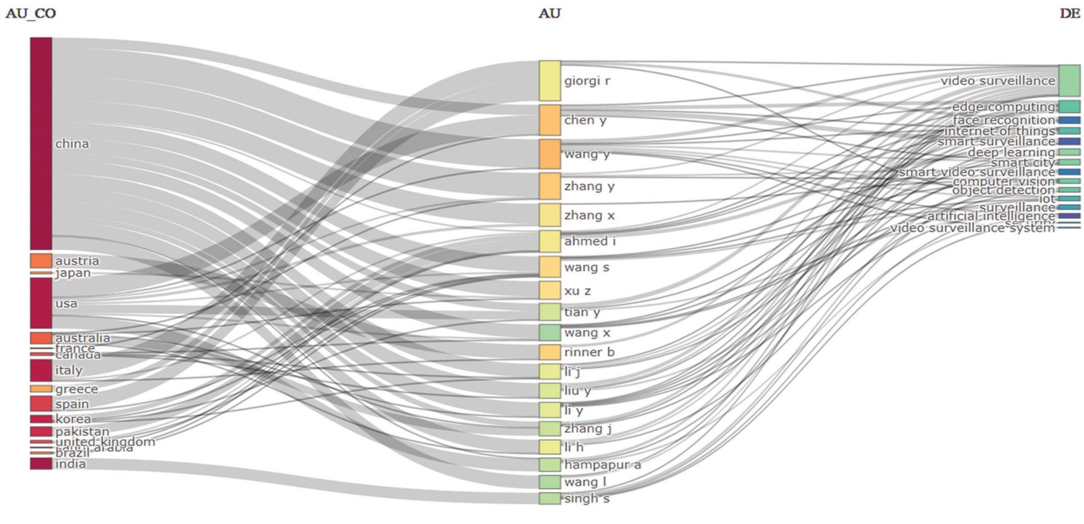


Fig. 5. Three-field chart (countries, authors, and keywords)

In Figure 5, the three-field graph, based on Sankey's diagrams [9], exposes the most influential authors and countries according to the most relevant keywords in the research.

It is noted that China is the country with the most publications, and the topic of "intelligent video surveillance" is widely discussed, especially among academics in the United States, Italy and Spain.

3.4. SCIENTIFIC CONTRIBUTION BY COUNTRY

In the research on "smart video surveillance", China is positioned as the main contributor with 535 publications. India follows with 402, and the United States has 199.

Italy, South Korea, Pakistan, Saudi Arabia, Spain, Germany and the United Kingdom complete the group of the 10 most active countries in this field, with 142, 104, 73, 63, 60, 59 and 57 publications, respectively (See Fig. 6).

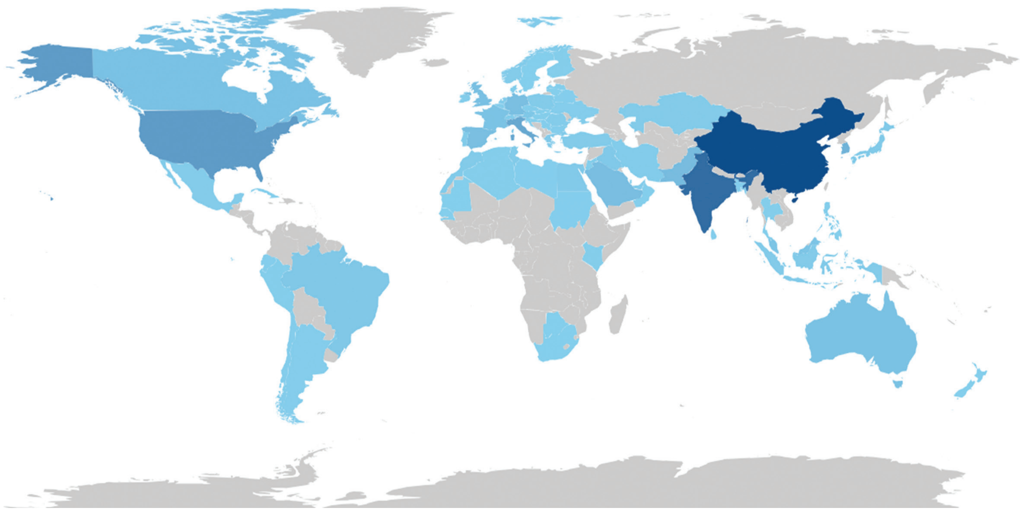


Fig. 6. Number of Articles by Country (blue shade: country or region with posts, grey shade: country or region with no posts, blue intensity: most posts)

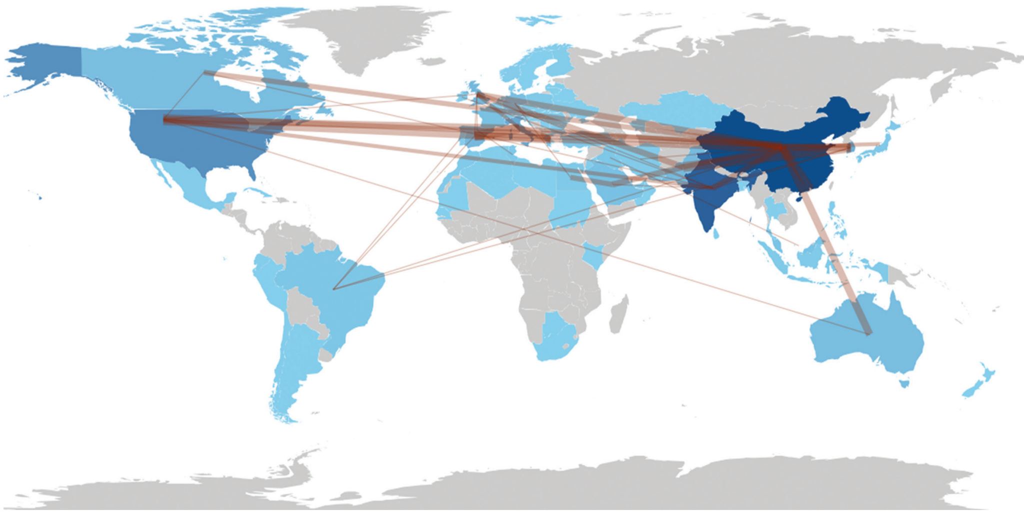


Fig. 7. Country collaboration map

3.5. MOST INTERNATIONALLY REFERENCED PUBLICATIONS

Table 1 shows the findings of the 20 most cited articles globally on Intelligent Video Surveillance [10-29], including CHAABOUNI N, 2019, IEEE COMMUN SURV TUTOR leads with 635 citations in total (105.83 per year), HAMPAPUR A, 2005, IEEE SIGNAL PROCESS MAG with 268 citations in total (13.40 per year); MEMOS VA, 2018, FUTURE GENER COMPUT SYST with 266 citations in total (38 per year); MINOLI D, 2018, INTERNET THING with 254 citations in total (36.29 per year); BAI Y, 2018, IEEE TRANS MULTIMEDIA with 234 citations in total (33.43 per year); ZHOU X, 2021, IEEE INTERNET THINGS J with 230 citations in total (57.50 per year); ZHANG T, 2015, PROC ANNU INT CONF MOBILE COMPUT NETWORKING with 227 citations in total (22.70 per year); The table includes articles published in a variety of relevant journals and conferences, not only in surveillance and security, but also in related fields such as communication networks, artificial intelligence, and mobile computing (e.g., Hampapur et al., 2005 in IEEE Signal Processing

Magazine and Minoli, 2018 in Internet of Things). This diversity shows that research on intelligent video surveillance is not limited to a single domain, but spans several disciplines, reinforcing its multidisciplinarity and applicability in different technological fields.

Table 1. Ranking of the 20 most referenced scientific articles worldwide

Paper	DOI	Total Citations
CHAABOUNI N, 2019, IEEE COMMUN SURV TUTOR	10.1109/COMST.2019.2896380	635
HAMPAPUR A, 2005, IEEE SIGNAL PROCESS MAG	10.1109/MSP.2005.1406476	268
MEMOS VA, 2018, FUTURE GENER COMPUT SYST	10.1016/j.future.2017.04.039	266
MINOLI D, 2018, INTERNET THING	10.1016/j.iot.2018.05.002	254
BAI Y, 2018, IEEE TRANS MULTIMEDIA	10.1109/TMM.2018.2796240	234
ZHOU X, 2021, IEEE INTERNET THINGS J	10.1109/JIOT.2021.3077449	230
ZHANG T, 2015, PROC ANNU INT CONF MOBILE COMPUT NETWORKING	10.1145/2789168.2790123	227
MEHTA P, 2020, COMPUT COMMUN	10.1016/j.comcom.2020.01.023	205
FRAGA-LAMAS P, 2017, SENSORS	10.3390/s17061457	175
LINGXIAO H, 2019, PROC IEEE INT CONF COMPUT VISION	10.1109/ICCV.2019.00854	165
YADAV SK, 2021, KNOWL BASED SYST	10.1016/j.knosys.2021.106970	158
NAYAK R, 2021, IMAGE VISION COMPUT	10.1016/j.imavis.2020.104078	158
TIAN Y, 2011, IEEE TRANS SYST MAN CYBERN PT C APPL REV	10.1109/TSMCC.2010.2065803	158
FLECK S, 2008, PROC IEEE	10.1109/JPROC.2008.928765	141
SHORFUZZAMAN M, 2021, SUSTAINABLE CITIES SOC	10.1016/j.scs.2020.102582	135
SAPONARA S, 2021, J REAL-TIME IMAGE PROCESS	10.1007/s11554-020-01044-0	130
WANG F, 2020, IEEE ACCESS	10.1109/ACCESS.2020.2982411	130
HOLTE MB, 2012, IEEE J SEL TOP SIGN PROCES	10.1109/JSTSP.2012.2196975	123
PAISITKRIANGKRAI S, 2008, IEEE TRANS CIRCUITS SYST VIDEO TECHNOL	10.1109/TCSVT.2008.928213	119
KE R, 2021, IEEE TRANS INTELL TRANSP SYST	10.1109/TITS.2020.2984197	113

3.6. KEYWORD ANALYSIS

Analyzing keywords is essential, as it allows us to identify the most relevant topics within the field of research [3]. Figure 8 reveals that "security systems" and "video surveillance" are the terms that predominate, evidencing that the central focus of research on intelligent video surveillance is on security and monitoring. This fact highlights the growing importance of surveillance systems in the protection of critical infrastructures, the prevention of crime and the reinforcement of both public and private security, which suggests a constant advance in the development and application of these technologies in different sectors.

Fig. 8. Word Cloud

Fig. 9 reflects the growth of the most significant keywords over the years. The top ten keywords comprise SECURITY SYSTEMS, VIDEO SURVEILLANCE, MONITORING, CAMERAS, DEEP LEARNING, VIDEO SURVEILLANCE SYSTEMS, NETWORK SECURITY, INTERNET OF THINGS, SMART CITY, and OBJECT DETECTION. Video surveillance analytics has advanced significantly with the use of deep learning and anomaly detection, which has allowed for more efficient identification of suspicious patterns in real-time. According to [30], these methods reduce the amount of data to be analyzed by focusing on critical patterns, alleviating the human workload. This advance explains the exponential growth of investigations into security and video surveillance systems observed in the graph since 2015.

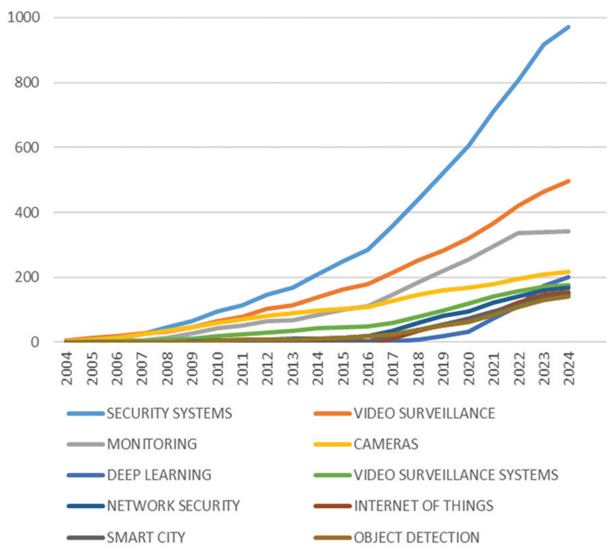


Fig. 9. Most Developed Words

3.7. MORE PRODUCTIVE JOURNALS

The thematic map provides a detailed interpretation of the patterns and trends of the topics of study, including information on seasonality and outliers. These maps divide the topics into four quadrants, using centrality on the X-axis and density on the Y-axis as classification criteria. Centrality indicates the connectivity and relevance of a topic in the general scope, while density reflects the degree of cohesion and development within the cluster [31]. As can be seen in Figure 10, the topics are classified into four categories: motors, the most connected and developed; basic topics, which although well connected, still have room to develop; niche topics, highly cohesive but less globally connected; and emerging issues, which are still nascent but with a potential significant future impact.

Motor themes. This quadrant includes the topics with the highest centrality and density, which means that they are well-developed topics with great impact within the field of intelligent video surveillance. Examples of these topics include "security," "real-time surveillance," and "object detection," which are critical in smart video surveillance applications. A clear example of a driving theme is CrowdSurge's article: A Crowd Density Monitoring Solution Using Smart Video Surveillance with Security Vulnerability Assessment. This article proposes a crowd density monitoring solution that uses intelligent video surveillance to manage public safety. This issue is critical in mass events or demonstrations, where security depends largely on the ability to monitor the density of people in real time, something that is also key in emergency management. This type of technology is developed in environments where safety is a priority concern [32]

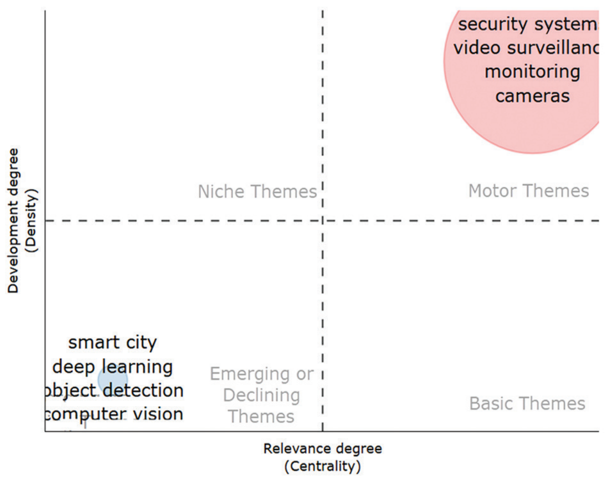


Fig. 10. Theme map

The basic themes. This quadrant includes topics that, although highly relevant, have not yet reached a high degree of development. These are central areas for the field that need more attention and research. A clear example of a basic topic in this quadrant is the article Abnormal human behavior detection in videos: A review. The detection of abnormal behaviors in surveillance systems is of great relevance, especially in public safety and crime prevention. However, despite their relevance, the development of practical solutions remains limited, with challenges in accuracy and real-time contextual recognition [5].

The Niche Themes: This quadrant is empty, indicating the absence of topics that are highly developed yet hold low central relevance to global research on intelligent video surveillance. In other words, there are no specialized areas within the field that are highly advanced but lack significant impact or importance in the broader context of the discipline.

Emerging or Declining Themes: This quadrant highlights various methods, techniques, and algorithms leveraging both online and offline deep learning. Identification algorithms, such as YOLO, are particularly effective in surveillance systems deployed in controlled environments like hallways, rooms, warehouses, banks, and parking lots. Additionally, highly accurate and efficient algorithms such as Faster-RCNN and RFCN are notable for detecting potential threats in images, such as identifying dangerous objects like weapons during baggage scans at airports. Offline algorithms are best suited for preprocessing tasks, such as analyzing an individual's behavior in previously recorded video sequences [26, 27]. Meanwhile, Sort and Deep Sort algorithms stand out as the most efficient for object tracking, balancing accuracy and speed, for example tracking pedestrians and vehicles.

4. CURRENT CHALLENGES AND FUTURE STUDIES

Intelligent Video Surveillance faces several key challenges related to the integration of emerging technologies such as IoT and artificial intelligence, which are still in the early stages of adoption and present interoperability problems between hybrid systems [10, 11]. In addition, cybersecurity and data privacy are critical concerns, due to the vulnerabilities of cloud-connected systems and the large amount of sensitive information that is collected [10, 12]. There are also challenges in real-time image processing and accurate object detection in dynamic environments, such as vehicular traffic, which requires high accuracy and speed [11, 14]. Latency and efficiency in managing large volumes of data in the cloud remain major issues, especially in connected cities, where the integration of multiple data sources is complex [12, 13]. Finally, energy consumption in devices with artificial intelligence is a key challenge to ensure continuous and efficient surveillance [14]. In summary, the main challenges encompass security, privacy, interoperability, and efficiency in the processing and energy consumption of these systems.

Smart Video Surveillance faces a number of key challenges that span several technological aspects. One of them is the synchronization between sensor networks and real-time processing algorithms, which requires advances in 5G [15], along with the heterogeneity of connected devices, which generates problems of standardization and compatibility of hardware and software. Compressing video to reduce bandwidth without affecting quality remains a challenge, especially in mobile applications, where resources are limited [16]. Quality of service (QoS) in video surveillance networks is another challenge, especially in areas with poor network infrastructure, which affects real-time transmission [17]. It is also necessary to improve the computational efficiency of machine learning models to handle large volumes of data and improve the accuracy in the detection of events and objects [33, 34]. In addition, latency in data transmission and scalability issues of IoT-based systems complicate real-time monitoring [18]. Finally, adverse lighting and weather conditions, as well as efficiency in training deep learning models with large datasets, represent significant computational challenges [19].

One of the great challenges of Intelligent Video Surveillance is the protection of privacy, as the vast amount of personal data handled raises ongoing concerns [20]. Detecting and tracking objects in dense urban environments, with multiple people and moving vehicles, remains challenging, especially in conditions of variable lighting and visual interference [35, 36]. In addition, the integration of fixed and mobile surveillance cameras poses data processing and compatibility issues, while synchronization between multiple cameras distributed in a system remains a considerable challenge [22, 23]. In addition, in sustainable cities, the high energy consumption of cameras and sensors poses challenges in terms of efficient energy use, while integration with smart urban technologies, such as transport networks and waste management, requires advances in interoperability and technological standards [24].

Future studies on smart video surveillance focus on incorporating new technologies, such as deep learning and IoT, to improve object segmentation and detection in dynamic and complex urban environments [21]. A key area of research is the refinement of multi-target tracking algorithms in distributed systems, along with the reduction of energy consumption for implementation in smart cities [37-39]. In addition, progress is needed in the synchronization and transmission of data between networked cameras, improving the efficiency of storage and processing in systems with limited infrastructures [40, 41]. Since smart cities require sustainability, it is crucial to develop energy-efficient technologies that do not compromise the performance of video surveillance systems [24]. Finally, improving real-time processing algorithms to handle large volumes of data without losing accuracy or speed is a major challenge in future research [25].

The future of intelligent video surveillance should focus on improving cybersecurity, developing more robust network architectures to protect data against cyberattacks [42, 43]. Another crucial aspect is the advance in the recognition of human actions through computer vision algorithms, which make it possible to distinguish between suspicious and routine behaviors in crowded environments such as airports and stations [44, 45]. In addition, research is needed to improve the compression and efficient transmission of high-resolution video, ensuring the quality of images in real time without consuming too much bandwidth, especially in areas with limited infrastructure [28, 29]. In the field of transport, studies should advance in the integration of video surveillance with road infrastructures to analyse the behaviour of drivers and pedestrians, improving road safety [46-48]. Finally, emerging technologies such as artificial intelligence and blockchain are essential to ensure data security and integrity, facilitating more reliable monitoring in complex scenarios such as smart cities [49-51].

ACKNOWLEDGEMENT

The authors express their gratitude to the Universidad Nacional de Cañete for supporting this work through Project 007-2023.

FUNDING

The research was funded by the Vice-Presidency of Research of the Universidad Nacional de Cañete (Project 007-2023).

5. REFERENCES:

[1] M. F. Nurnoby, T. Helmy, "A real-time deep learning-based smart surveillance using fog computing: a complete architecture", Proceedings of the International Conference on Machine Learning and Data Engineering, Dehradun, Saudi Arabia, 7-8 September 2022, pp. 1102-1111.

[2] T. Golda, D. Guaia, V. Wagner-Hartl, "Perception of Risks and Usefulness of Smart Video Surveillance Systems", Applied Sciences, Vol. 12, No. 20, 2022, p. 10435.

[3] Y. Chen, L. Shu, L. Wang, "Poster abstract: Traffic flow prediction with big data: A deep learning based time series model", Proceedings of the IEEE Conference on Computer Communications Workshops, Atlanta, GO, USA, 1-4 May 2017, pp. 1010-1011.

[4] G. Aldehim, M. M. Asiri, M. Aljebreen, A. Mohamed, M. Assiri, S. S. Ibrahim, "Tuna Swarm Algorithm With Deep Learning Enabled Violence Detection in Smart Video Surveillance Systems", IEEE Access, Vol. 11, 2023, pp. 95104-95113.

[5] H. Mu, R. Sun, G. Yuan, Y. Wang, "Abnormal Human Behavior Detection in Videos: A Review", Information Technology and Control, Vol. 50, No. 3, 2021, pp. 522-545.

[6] P. Mongeon, A. Paul-Hus, "The journal coverage of Web of Science and Scopus: a comparative analysis", Scientometrics, Vol. 106, No. 1, 2016, pp. 213-228.

[7] G. Halevi, H. Moed, J. Bar-Ilan, “Suitability of Google Scholar as a source of scientific information and as a source of data for scientific evaluation-Review of the Literature”. Journal of Informetrics, Vol. 11, No. 3, 2017, pp. 823-834.

[8] C. Dwork, A. Roth, “The algorithmic foundations of differential privacy”, Foundations and Trends in Theoretical Computer Science, Vol. 9, No. 3-4, 2013, pp. 211-407.

[9] P. Donner, U. Schmoch, “The implicit preference of bibliometrics for basic research”, Scientometrics, Vol. 124, No. 2, 2020, pp. 1411-1419.

[10] N. Chaabouni, M. Mosbah, A. Zemmari, C. Sauvignac, P. Faruki, “Network Intrusion Detection for IoT Security Based on Learning Techniques” IEEE Communications Surveys and Tutorials, Vol. 21, No. 3, 2019, pp. 2671-2701.

[11] A. Hampapur, L. Brown, J. Connell, A. Ekin, N. Haas, M. Lu, H. Merkl, S. Pankanti, “Smart video surveillance: exploring the concept of multiscale spatiotemporal tracking”, IEEE Signal Processing Magazine, Vol. 22, No.2, 2005, pp. 38-51.

[12] V. A. Memos, K. E. Psannis, Y. Ishibashi, B. Kim, B. B. Gupta, “An Efficient Algorithm for Media-based Surveillance System (EAMSuS) in IoT Smart City Framework”, Future Generation Computer Systems, Vol. 83, 2018, pp. 619-628.

[13] D. Minoli, B. Occhiogrosso, “Blockchain mechanisms for IoT security”, Internet of Things, Vol. 1-2, 2018, pp. 1-13.

[14] Y. Bai, Y. Lou, F. Gao, S. Wang, Y. Wu, L. Duan, “Group-Sensitive Triplet Embedding for Vehicle Reidentification”, IEEE Transactions on Multimedia, Vol. 20, No. 9, 2018, pp. 2385-2399.

[15] X. Zhou, X. Xu, W. Liang, Z. Zeng, Z. Yan, “Deep-Learning-Enhanced Multitarget Detection for End-Edge-Cloud Surveillance in Smart IoT”, IEEE Internet of Things Journal, Vol. 8, No. 16, 2021, pp. 12588-12596.

[16] T. Zhang, A. Chowdhery, P. Bahl, K. Jamieson, S. Banerjee, “The design and implementation of a wireless video surveillance system”, Proceedings of the Annual International Conference on Mobile Computing and Networking, MOBICOM, Paris, France, 7-11 September 2015, pp. 426-438.

[17] P. Mehta, R. Gupta, S. Tanwar, “Blockchain envisioned UAV networks: Challenges, solutions, and comparisons”, Computer Communications, Vol. 151, 2020, pp. 518-538.

[18] P. Fraga-Lamas, T. M. Fernández-Caramés, L. Castedo, “Towards the Internet of Smart Trains: A Review on Industrial IoT-Connected Railways”, Sensors, Vol. 17, No. 6, 2017, p. 1457.

[19] H. Lingxiao, Y. Wang, W. Liu, H. Zhao, Z. Sun, J. Feng, “Foreground-Aware Pyramid Reconstruction for Alignment-Free Occluded Person Re-Identification”, Proceedings of the 17th IEEE/CVF International Conference on Computer Vision, Seoul, Korea, 27 October - 2 November, pp. 8449-8458.

[20] S. K. Yadav, K. Tiwari, H. M. Pandey, S. A. Akbar, “A review of multimodal human activity recognition with special emphasis on classification, applications, challenges and future directions”, Knowledge-Based Systems, Vol. 223, 2021, p. 106970.

[21] R. Nayak, U. C. Pati, S. K. Das, “A comprehensive review on deep learning-based methods for video anomaly detection”, Image and Vision Computing, Vol. 106, 2021, p. 104078.

[22] Y. Tian, R. S. Feris, H. Liu, A. Hampapur, M. Sun, “Robust Detection of Abandoned and Removed Objects in Complex Surveillance Videos”, IEEE Transactions on Systems, Man and Cybernetics Part C: Applications and Reviews, Vol. 41, No. 5, 2011, pp. 565-576.

[23] S. Fleck, W. Straßer, “Smart Camera Based Monitoring System and Its Application to Assisted Living”, Proceedings of the IEEE, Vol. 96, No. 10, 2008, pp. 1698-1714.

[24] M. Shorfuzzaman, M. S. Hossain, M. F. Alhamid, “Towards the sustainable development of smart cities through mass video surveillance: A response to the COVID-19 pandemic”, Sustainable Cities and Society, Vol. 64, 2021, p. 102582.

[25] S. Saponara, A. Elhanashi, A. Gagliardi, “Real-time video fire/smoke detection based on CNN in antifire surveillance systems”, Journal of Real-Time Image Processing, Vol. 18, No. 3, 2021, pp. 889-900.

[26] F. Wang, M. Zhang, X. Wang, X. Ma J. Liu, “Deep Learning for Edge Computing Applications: A State-of-the-Art Survey”, IEEE Access, Vol. 8, 2020, pp. 58322-58336.

[27] D. Ameijeiras, H. R. González, Y. Hernández, "Algorithms for detection and tracking objects with deep networks for intelligent video surveillance: A review", Revista Cubana de Ciencias Informáticas, Vol. 14, No. 3, 2020, pp. 165-195.

[28] K. U. Duja, I. A. Khan, M. Alsuhaibani, "Video Surveillance Anomaly Detection: A Review on Deep Learning Benchmarks", IEEE Access, Vol. 12, 2024, pp. 164811-164842.

[29] R. Ke, Y. Zhuang, Z. Pu Y. Wang, "A Smart, Efficient, and Reliable Parking Surveillance System With Edge Artificial Intelligence on IoT Devices", IEEE Transactions on Intelligent Transportation Systems, Vol. 22, No. 8, 2021, pp. 4962-4974.

[30] B. Asal, A. B. Can, "Ensemble-Based Knowledge Distillation for Video Anomaly Detection", Applied Sciences, Vol. 14, No. 3, 2024, p. 1032.

[31] M. M. Najafabadi, F. Villanustre, T. M. Khoshgoftaar, N. Seliya, R. Wald, E. Muharemagic, "Deep learning applications and challenges in big data analytics", Journal of Big Data, Vol. 2, No. 1, 2015, pp. 1-21.

[32] M. J. Samonte, A. C. Garcia, J. E. Gorre, J. A. K. Perez, "CrowdSurge: A Crowd Density Monitoring Solution Using Smart Video Surveillance with Security Vulnerability Assessment", Journal of Advances in Information Technology, Vol. 13, No. 12, 2022, pp. 173-180.

[33] S. Wu, Q. Huang, L. Zhao, "A deep learning-based network for the simulation of airborne electromagnetic responses", Geophysical Journal International, Vol. 233, No. 1, 2023, pp. 253-263.

[34] S. Stockman, D. Lawson, M. Werner, "SB-ETAS: using simulation based inference for scalable, likelihood-free inference for the ETAS model of earthquake occurrences", Statistics and Computing, Vol. 34, No. 5, 2024, p. 174.

[35] Y. Tang, J. Zhang, R. Liu, Y. Li, "Exploring the Impact of Built Environment Attributes on Social Followings Using Social Media Data and Deep Learning", ISPRS International Journal of Geo-Information, Vol. 11, No. 6, 2022, p. 325.

[36] F. Zhang, Y. Liu, "Street view imagery: Methods and applications based on artificial intelligence", National Remote Sensing Bulletin, Vol. 25, No. 5, 2021, pp. 1043-1054.

[37] M. N. Alatawi, "Optimization of Home Energy Management Systems in Smart Cities Using Bacterial Foraging Algorithm and Deep Reinforcement Learning for Enhanced Renewable Energy Integration", International Transactions on Electrical Energy Systems, Vol. 2024, 2024, p. 2194986.

[38] B. Kommey, E. Tamakloe, J. J. Kponyo, E. T. Tchao, A. S. Agbemenu, H. Nunoo-Mensah, "An artificial intelligence-based non-intrusive load monitoring of energy consumption in an electrical energy system using a modified K-Nearest Neighbour algorithm", IET Smart Cities, Vol. 6, No. 3, 2024, pp. 132-155.

[39] A. Pollak, A. Gupta, D. Göhlich, "Optimized Operation Management With Predicted Filling Levels of the Litter Bins for a Fleet of Autonomous Urban Service Robots", IEEE Access, Vol. 12, 2024, pp. 7689-7703.

[40] Z. Peng, J. Li, H. Hao, Y. Zhong, "Smart structural health monitoring using computer vision and edge computing", Engineering Structures, Vol. 319, 2024, p. 118809.

[41] Y. Shao, L. Li, J. Li, Q. Li, S. An, H. Hao, "3D displacement measurement using a single-camera and mesh deformation neural network", Engineering Structures, Vol. 318, 2024, p. 118767.

[42] E. Naderi, A. Asrari, B. Ramos, "Moving Target Defense Strategy to Protect a PV/Wind Lab-Scale Microgrid Against False Data Injection Cyberattacks: Experimental Validation", Proceedings of the IEEE Power and Energy Society General Meeting, Orlando, FL, USA, 16-20 July 2023, pp. 1-5.

[43] E. Naderi, A. Asrari, "Detection of False Data Injection Cyberattacks: Experimental Validation on a Lab-scale Microgrid", Proceedings of the IEEE Green Energy and Smart Systems, Long Beach, CA, USA, 7-8 November 2022, pp. 1-6.

[44] P. Pereira, J. Araujo, C. Melo, V. Santos, P. Maciel, "Analytical models for availability evaluation of edge and fog computing nodes", Journal of Supercomputing, Vol. 77, 2021, pp. 9905-9933.

[45] P. Pereira, C. Melo, J. Araujo, J. Dantas, V. Santos, P. Maciel, "Availability model for edge-fog-cloud continuum: an evaluation of an end-to-end infrastructure of intelligent traffic management service", Journal of Supercomputing, Vol. 78, 2022, pp. 4421-4448.

[46] R. Liu, M. Li, L. Ma, "Efficient in-situ image and video compression through probabilistic image representation", Signal Processing, Vol. 215, 2024, p. 109268.

[47] M. K. Sharma, I. Farhat, C. Liu, N. Sehad, W. Hamidouche, M. Debbah, "Real-Time Immersive Aerial Video Streaming: A Comprehensive Survey, Benchmarking, and Open Challenges", IEEE Open Journal of the Communications Society, Vol. 5, 2024, pp. 5680-5705.

[48] C. Chen, C. Wang, B. Liu, C. He, L. Cong, S. Wan, "Edge Intelligence Empowered Vehicle Detection and Image Segmentation for Autonomous Vehicles", IEEE Transactions on Intelligent Transportation Systems, Vol. 24, No. 11, 2023, pp. 13023-13034.

[49] A. Kumari, S. Tanwar, "A secure data analytics scheme for multimedia communication in a decentralized smart grid", Multimedia Tools and Applications, Vol. 81, No. 24, 2022, pp. 34797-34822.

[50] S. J. Patil, L. S. Admuthe, A. S. Patil, S. R. Prasad, "Secure MANET routing with blockchain-enhanced latent encoder coupled GANs and BEPO optimization", Smart Science, Vol. 12, No. 4, 2024, pp. 1-14.

[51] H. A. Ahmed, H. A. A. AL-Asadi, "An Optimized Link State Routing Protocol with a Blockchain Framework for Efficient Video-Packet Transmission and Security over Mobile Ad-Hoc Networks", Journal of Sensor and Actuator Networks, Vol. 13, No. 2, 2024, p. 22.

DCT-based Robust Reversible Watermarking Technique based on histogram Modification

Original Scientific Paper

Soumitra Roy
Department of Computer Science and Engineering,
Dr. Sudhir Chandra Sur Institute of Technology & Sports Complex, Kolkata, West Bengal - 700074, India
mosinapur.sou@gmail.com

Naushad Varish*
Department of Computer Science and Engineering,
GITAM (Deemed to be University), Hyderabad Campus
Sangareddy-502329, Telangana, India
naushad.cs88@gmail.com

Syed Irfan Yaqoob
Department of computer science Engineering -Apex institute of Technology Chandigarh University, Gharuan, Mohali, Punjab, 140413, India.
syedirfan.ssm@gmail.com

*Corresponding author

Md Shamsul Haque Ansari
Department of Computer Science and Engineering,
Koneru Lakshmaiah Education Foundation,
Vaddeswaram, Guntur-522302, A.P. India
shamsshamsul@gmail.com

Abu Taha Zamani
Department of Computer Science, Faculty of Science,
Northern Border University,
Arar, Kingdom of Saudi Arabia
abutaha.zamani@nbu.edu.sa

Abstract *– In this paper, a strong, reversible image watermarking technique based on discrete cosine transform (DCT) and histogram shifting is proposed, where it overcomes the following concerns: (i) Reversing the cover object to its starting appearance is the primary goal of the reversible watermarking system. (ii) Military, medical, and standard law enforcement images are the main types of images that require distortion and reinstatement of the cover object following the watermark extraction. (iii) Lack of robustness and cover image-dependent embedding capacity are the primary concerns about reversible watermarking. Decompose the cover object into blocks that don't overlap in the first stage to insert a binary watermark bit into every block that is converted. These binary bits of watermark are embedded by altering a single set of middle substantial AC coefficients. To restore the cover image, subsequently using the histogram bin shifting method, a location map is created and integrated within the cover image. On the extracting side, at first, a location map is extracted from the image using the histogram bin shifting technique. In the following step, the image's watermark is recovered, and a reversed image has been generated using a location map. To verify the robustness property, several image processing attacks are tested with the suggested reversible watermarking approach, and favorable results are attained. The proposed scheme using the Lena image achieved 46.62 imperceptibility for 4096 embedding capacities. To methodically evaluate the proposed approach, it is compared with two current reversible watermarking systems, where they achieved 39.10 and 37.90 imperceptibility with 4.4 × 103 and 256 embedding capacities, respectively. The experimental results affirmed that the suggested method exhibits superior performance relative to these existing techniques.*

Keywords: *Reversible watermarking; DCT; Robustness; Histogram modification*

Received: November 23, 2024; Received in revised form: March 13, 2025; Accepted: March 14, 2025

1. INTRODUCTION

Recently, scholars have innovated digital watermarking as a comprehensive methodology to furnish intrinsic security [1] for digital data. According to Mintzer et al. [2], there exist three distinct categories of watermarking applications:

Ownership Assertion – Structured to convey ownership information. Integrity Verification – Guarantees that the content of the item remains unaltered. Captioning – Delivers object-specific information or annotations to a designated community of users.

Mintzer et al. assert that while ownership and captioning watermarks exhibit robustness, watermarks predicated on content alteration are inherently fragile. An alternative category of watermarking, referred to as semi-fragile watermarking, possesses the capacity to withstand certain forms of attacks.

Identification codes and watermarks, particularly those that incorporate the proprietor's details or corporate insignias, are irrevocably inscribed in the cover object for subsequent verification utilizing watermarking methodologies. Nevertheless, when the cover image is irretrievably modified, a significant concern arises regarding the potential loss of essential information embedded in the cover.

In domains such as law enforcement, military operations, and medical practices, the paramount requirement is the lossless or distortion-free restoration of the cover object after watermark removal. These contexts typically utilize lossless or reversible watermarking methodologies.

Existing reversible watermarking methods are grouped into five main classes: (i) compression domain, (ii) transform domain, (iii) quantization-based, (iv) integration of encryption and data concealing, and (v) spatial domain. In all these reversible watermarking algorithms, data hiding space is generated in the cover image for watermark embedding. As a result, it is very hard to put a distinct boundary among several classes of procedures for reversible watermarking. The subsequent segment provides a brief overview of various kinds of reversible watermarking methods. In the next section, brief overviews of various reversible watermarking methodology types are discussed.

Compressed domain methods: The compressed sector reversible approaches include techniques such as (i) vector quantization [3-5], (ii) block truncation coding [6], (iii) MPEG coding [7], and (iv) least significant bit (LSB) insertion employing data compression [8-10].

Quantization-based: The methods of watermarking that rely on reversible quantization are fragile. However, watermarking techniques based on quantization are generally reliable. This reversible watermarking cluster includes the algorithms presented in [11-13].

Transform domain: Transform domain reversible watermarking algorithms are based on (i) Integer DCT [14, 15] and (ii) Integer Wavelet Transform (IWT) [16].

Combination of data hiding and encryption: A few innovative reversible data hiding techniques that combine encryption and data concealing are shown in [17, 18].

Spatial domain:

(i) Difference expansion (DE): Expanding the converted integers is how watermark bits are added in DE-based reversible watermarking techniques. Based on its underlying concepts, DE-based algorithms can be divided into five classes: i) General Difference Expansion [19-31], ii) Companding technique [32], iii) Contrast mapping [33, 34], iv) Prediction error based [35-51], and v) Interpolation [52, 53].

(ii) Histogram modification: Histogram bin shifting approaches inject the watermark by using the image's histogram. These types of algorithms are presented in [54-66].

Main contribution of work: This research proposes a novel digital watermarking approach based on DCT and histogram modification that is robust and reversible. The rationale for using DCT to implement this reversible image watermarking technique, as well as certain unique aspects of the suggested approach, are explained in depth in the section that follows:

- The "blocking artifact," which arises at block boundaries because of imprecise quantization of the coefficients, is one of the main drawbacks of block-based DCT. The suggested DCT-based reversible watermarking system is created by employing a single pair of middle spectrum components after row-major scanning for each DCT block to minimize the problem of visual artifacts. Compared to low and high coefficient pairings, these medium band coefficient pairs are less susceptible to alteration.
- A simple procedure is used in the presented blind image watermarking methodology to maintain one watermark bit in each fragmented non-overlapping parent image block. Following row-major scanning, choose one pair of DCT coefficients at a time from the middle bands. Next, determine if the first coefficient is bigger than the second coefficient for every coefficient couple to maintain the watermark bit=1. Retain the same coefficients if the condition is met. If not, switch these two values. In each pair of coefficients, the first coefficient must be less than the second to maintain the watermark bit=0. Repetitive bits are added to the host image in traditional error-correcting code (ECC)- oriented watermarking systems to aid in error correction or detection on the receiver side. Therefore, compared to previous ECC-based watermarking techniques, the suggested reversible watermarking scheme's implementation is simpler and has less computing complexity.

The remainder of the article is structured as described after this opening part. In section 2, the block-based DCT preliminary results are shown. In section 3, the suggested watermark extraction and embedding algorithms are explained. Section 4 provides experimental outcomes. Conclusions are finally stated in section 5.

2. BLOCK-BASED DCT

One of the most well-liked and frequently applied signal compression and decomposition methods is the DCT, which converts a signal from a spatial domain form into a spectrum demonstration with the intrinsic capacity to show superior energy compaction of the signal or picture. In essence, it changes the signal into an accumulation of sinusoids with different frequencies and magnitudes. The DCT conversion is used to move an image's pixel values from a particular domain to another; the resultant image has several AC coefficients and one DC coefficient. When using block-based DCT, a host image with dimensions of $M \times N$ is divided into non-overlapping blocks with dimensions of $m \times n$. Each block, denoted as fb, is then converted into a matching DCT coefficient using the equation that follows:

$$F_b(u,v) = \alpha(u)\alpha(v)\sum_{x=0}^{m-1}\sum_{y=0}^{n-1} f_b(x,y)\cos\left[\frac{(2x+1)u\pi}{2m}\right]\cos\left[\frac{(2y+1)v\pi}{2n}\right] \quad (1)$$

where,

$$\alpha(\mathrm{u}) = \begin{cases} \sqrt{1/m}, & u=0 \\ \sqrt{2/m}, & otherwise \end{cases}$$

$$\alpha(v) = \begin{cases} \sqrt{1/n}, & v=0 \\ \sqrt{2/n}, & otherwise \end{cases} \quad (2)$$

After the sub-image block $Fb(u, v)$ is modified, the sub-image is rebuilt by using.

$$f_b(x,y) = \alpha(u)\alpha(v)\sum_{u=0}^{m-1}\sum_{v=0}^{n-1} F_b(u,v)\cos\left[\frac{(2x+1)u\pi}{2m}\right]\cos\left[\frac{(2y+1)v\pi}{2n}\right] \quad (3)$$

for x=0, 1, 2, …, m-1, and y=0,1, 2,...,n-1 and $\propto$ is defined as in equation 2.

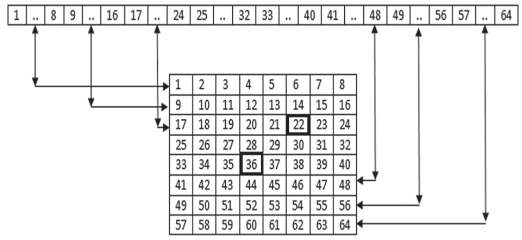


Fig. 1. Selected DCT coefficient pair of $m \times n$ image-block according to row-major scanning order

Three distinct frequency bands—the low, middle, and high-frequency bands—are produced by block-based DCT. Since the low-frequency band contains the most picture information, altering it generally distorts the image's perceived quality, whereas the high-frequency spectrum can be eliminated for compression purposes. For this reason, the middle band frequency is used in the development of DCT-based watermarking schemes because it is less noticeable when modified. The coefficient pair chosen for the watermark insertion in this suggested study is displayed in Fig. 1.

3. PROPOSED REVERSIBLE WATERMARKING SCHEME

This section provides a detailed explanation of the proposed reversible watermarking system based on DCT.

3.2. LOCATION MAP AND WATERMARK EMBEDDING PROCESS

Fig. 2 shows the procedure of inserting the location map and watermark.

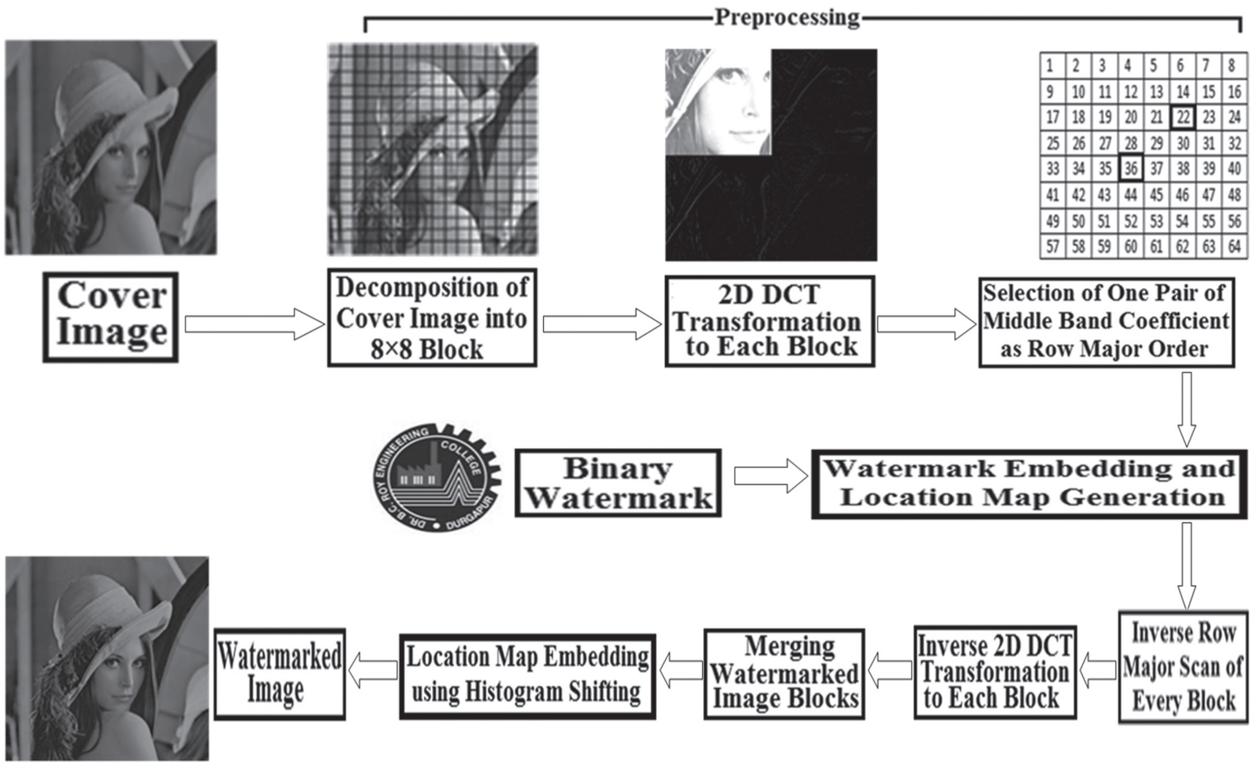


Fig. 2. Location map and watermark and location map implanting process

Algorithm 1: Watermark and Location map Embedding

Input: A cover image and a binary logo

Output: The watermarked image

Begin

Preprocessing for the implanting of the watermark: Preprocessing stages for this suggested blind image watermarking approach that embeds a binary watermark within a greyscale image are the ones that follow:

Step 1: Divide the greyscale host image, measuring M by N, into non-overlapping blocks, each measuring m by n.

Step 2: First, implement two-dimensional DCT at the block level in every non-overlapping block. In the following step, make 2D DCT coefficients into one-dimensional DCT factors by row-major order (as depicted in Figure). Then select two middle band coefficients to integrate. The values of this coefficient pair are kept secure as a secret code.

Watermark implanting and generating a location map: The suggested watermark implanting method is explained in the subsequent section below:

Step 3: The following guidelines are followed while embedding each bit of the binary logo:

Rule 1: When inserting a binary logo bit of 1: In the chosen pair of coefficients, verify if the first coefficient is smaller than the second, and then switch these two values. If not, do not alter the coefficients.

Rule 2: When inserting a binary logo bit of 0: Verify that the first coefficient in the chosen pair of coefficients is bigger than the second and then switch these two values. If not, maintain the same coefficients.

Step 4: For the first two rules, change the location map bit to 1 in the event of a swap, and set it to 0 otherwise.

After implanting the watermark, post-processing: After the binary logo is embedded, the following post-processing procedures are needed to obtain a watermarked image:

Step 5: After implementing the inverse DCT on each modified block, carry out the inverse DCT at the block level.

Step 6: Reconstructing the watermarked image involves combining all the altered blocks into a single block.

Embedding location map using histogram shifting: The following procedures are used to insert the location map created in step 4 into the cover image:

Step 7: The greyscale watermarked picture histogram $H_i=\{h_i \mid i=0,1,2,\ldots,255\}$ should be calculated, with h_i Denoting the ith bin's histogram value.

Step 8: Using the following formula, determine the histogram's peak (p) and minimum point (m):

$$\begin{cases} h_p = \max\{h_i\} \\ h_m = \min\{h_i\} \end{cases} \tag{4}$$

where, $\{i, p, m\} = 0,1,2,\ldots,255$.

Step 9: Using the following equation, analyze the watermarked object and accordingly adjust the pixel's intensity:

$$\begin{cases} i & if\ i \le p \\ i+1 & if\ p < i \le m \end{cases} \tag{5}$$

Step 10: Utilizing the following guidelines, revisit the watermarked image and integrate the location map:

$$i = \begin{cases} i & if\ i = p\ and\ location\ map\ bit = 0 \\ i = i+1 & if\ i = p\ and\ location\ map\ bit = 1 \\ i & otherwise \end{cases} \tag{6}$$

End

3.2. ORIGINAL COVER IMAGE RESTORATION POST WATERMARK EXTRACTION PROCESS

Fig. 3 illustrates the procedure of restoring the host after extracting the watermark

Algorithm 2: Watermark Extracting and Original Cover Image Restoring

Input: A modified/attack image

Output: A binary Logo and the reversible cover image

Begin

Retrieving the location map data:

Step 1: Utilizing the following equation, scan the altered/attack image and retrieve the location map segments:

$$location\ map\ bit = \begin{cases} 0\ if\ i' = p \\ 1\ if\ i' = p+1 \end{cases} \tag{7}$$

where p is the histogram's peak point, i' represents the updated image's pixel value, and $\{i, p\} = \{0,1, 2,\ldots,255\}$.

Preprocessing for watermark extraction: The following preprocessing procedures are included in this suggested blind image watermarking approach, which extracts a binary watermark from a greyscale image:

Step 2: Divide the greyscale host image, measuring M by N, into non-overlapping blocks, each measuring m by n.

Step 3: Choose two middle band factors based on major order for each non-overlapping block after applying block-level two-dimensional DCT. The coefficient pair values chosen here are those that are retained on the embedding side. In actuality, the values of these coefficient pairs originate from secret key data. As an additional payload, the watermarked image is transmitted along with secret key values.

Watermark extracting: The suggested watermark-extracting method is explained in the subsequent section below:

Step 4: Retrieve the single logo bit from each block using the following guidelines based on a pair of chosen DCT coefficients:

Rule 1: The codeword bit=1 if the first coefficient's intensity in the DCT pair is higher than or equal to the second coefficient's intensity.

Rule 2: The codeword bit=0 if the first coefficient's intensity in the DCT pair is lower than the second coefficient's intensity.

Step 5: The watermark bit streams that are retrieved and recovered from the altered cover picture should be stored. Reshape the retrieved watermark stream of bits into a 2-D matrix form to generate the watermark logo.

Obtaining a reversible cover image: The tasks listed below must be used to return the cover image to its original state:

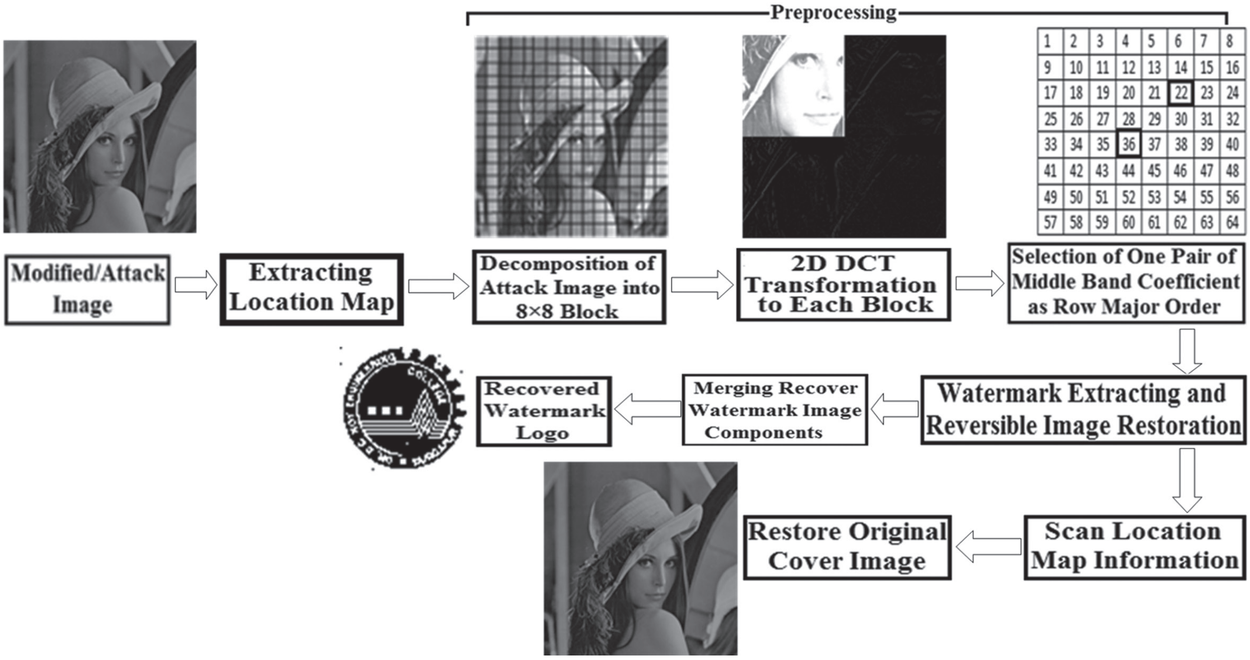


Fig. 3. Process for original host image restoration after watermark extraction

Step 6: The following rule is being used to restore each block of the cover image to its original shape along with the location map:

Rule: Examine the values of the coefficient pairs that were chosen in step three. If the matching location map bit is 1, swap the values of the specified DCT coefficient pair. Retain the chosen coefficient pair in case the matching location map bit value is 0.

Post-processing after watermark extraction: Here are the post-processing procedures to obtain the original image after removing the location map and binary logo:

Step 7: After implementing the inverse DCT on each modified block, carry out the inverse DCT at the block level.

Step 8: Reconstructing the reversible cover image by combining all the altered blocks into a single block.

End

4. EXPERIMENTAL RESULT

The described reversible watermarking method's effectiveness is tested against several experiments. The MATLAB platform is utilized for conducting these experiments on a selection of standard 512×512 images, including Peppers, Elain, Pirate, Zelda, Lena, Goldhill, Clown, Military (Indian Missile Agni-3), and Medical (malignant melanoma with bone marrow carcinomatosis) images. Additionally, a 64×64 logo is watermarked (as illustrated in Fig. 4).

This robust and reversible image watermarking method starts with applying an 8×8 block-based DCT to the source object. Then, from every converted block, a single middle-band AC component couple is chosen to implant the watermark based on the row-major scanning order. The efficacy of the suggested plan is demonstrated by comparing the suggested DCT-based reversible watermarking method to several trials in terms of (i) Reversibility, (ii) Imperceptibility, (iii) Robustness, and (iv) Embedding capacity.

Reversibility: Beyond retrieval of the watermark, the primary goal of the reversible watermarking technique is to return the host object to its initial formation. To assess the reversibility property of the suggested approach, a comparison is made between the bit error rate (BER) and the normalized cross-correlation (NC) value between the original cover image vs the reconstructed host object following watermark extraction. Table 1 displays the BER and NC values of a few standard images.

The NC value has a range of -1 to +1. This correlation value is around 1 if the restored cover image closely resembles the original, and -1 if it is negatively correlated with the cover image. If the NC value trends toward zero, it becomes completely unsatisfactory or irrelevant. Here is how to compute the BER and NC:

$$\text{BER}=\frac{\text{Number of error bits}}{\text{Total bits transmitted}}=\frac{\text{Number of error bits per second}}{\text{Data rate per second}} \tag{8}$$

$$\text{NC}(w,\bar{w})=\frac{\sum_{i=1}^{M}\sum_{j=1}^{N}[w(i,j)-\mu_w]\times[\bar{w}(i,j)-\mu_{\bar{w}}]}{\sqrt{\sum_{i=1}^{M}\sum_{j=1}^{N}[w(i,j)-\mu_w]^2}\times\sqrt{\sum_{i=1}^{M}\sum_{j=1}^{N}[\bar{w}(i,j)-\mu_{\bar{w}}]^2}} \tag{9}$$

For $M \times N$ dimension image: μw=mean of the host image, $\mu_{\bar{w}}$ = mean of the restored cover image; $w(i, j)$ =the pixel intensity value at coordinate (i, j) of the host object, $(\bar{w})(i, j)$=the pixel intensity value at coordinates (i, j) of the restored host object after watermark extraction respectively.

(ii) Imperceptibility Measurement: The change in perceptual picture quality caused by the suggested watermarking technique needs to be identified in order

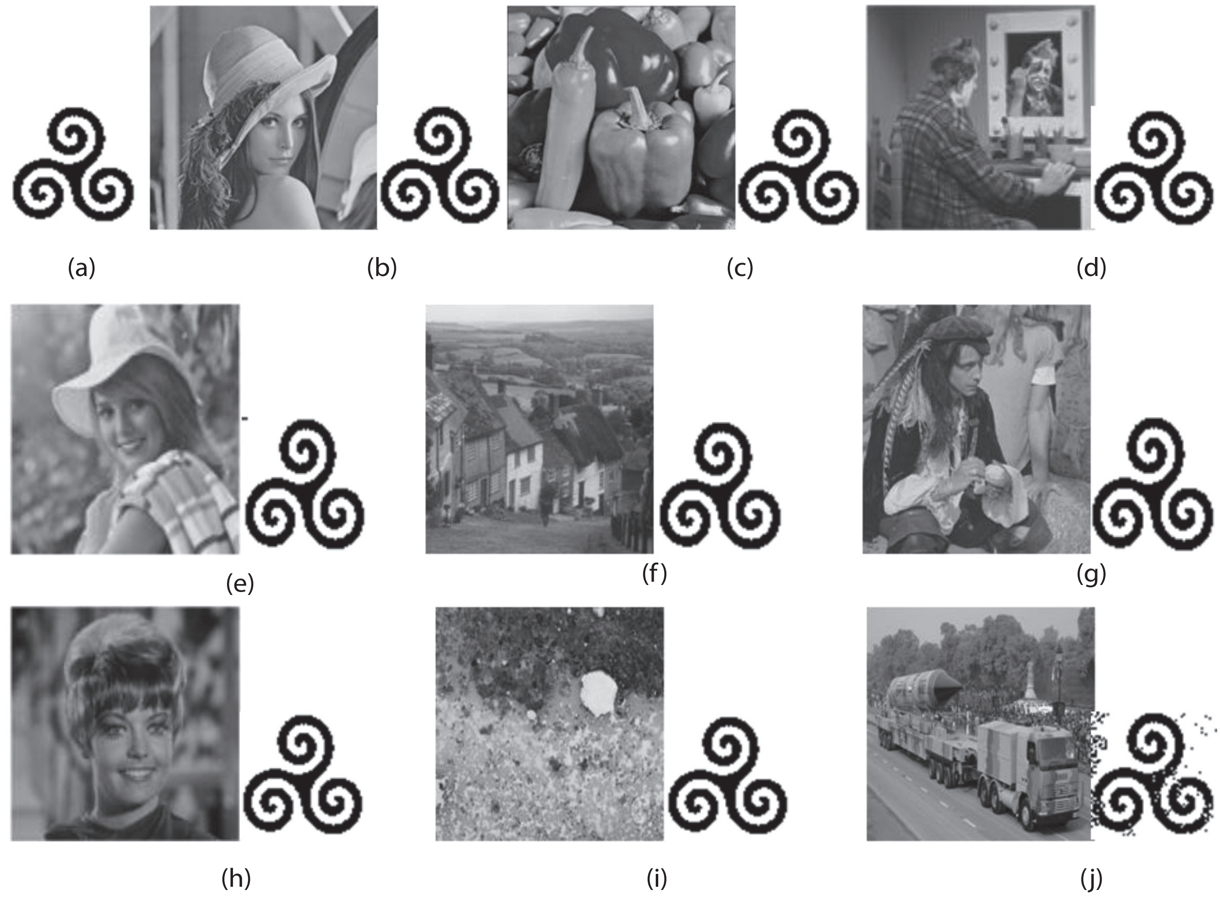


Fig. 4. (**a**) Initial watermark(binary),
(**b**)-(**j**)Recovered watermark with the corresponding images with no attack

A watermarking technique needs to be identified to compute imperceptibility/invisibility assessment. To determine the perceptual resemblance between a host image and the corresponding watermarked material, one uses the peak signal-to-noise ratio (PSNR). An efficient invisible watermarking technique should: (i) have a watermark that is undetectable or invisible to HVS and (ii) compare its results to a standard benchmark PSNR. The decibel (dB) represents the PSNR value. According to Petitcolas [67], 38 dB is the lowest permissible PSNR value for optimal imperceptibility. Since PSNR has no real significance when considering geometric distortions, this convention is questionable [68]. Here is how PSNR is defined:

$$\text{PSNR}=10\times\log_{10}\frac{\max(x(i,j))^2}{\text{MSE}} \tag{10}$$

In this case, the watermarked picture $\bar{x}$. The host image x is determined by their mean square error (MSE) as follows:

$$\text{MSE}=\frac{1}{M\times N}\sum_{i=1}^{M}\sum_{j=1}^{N}(x_{ij}-\bar{x}_{ij})^2 \tag{11}$$

In this case, the symbols M and N stand for the image's width and height, $\bar{x}$ for the initial image's pixel intensity measurement at coordinates (i, j), and $\bar{x}_{ij}$ for the watermarked image's corresponding value. In essence, PSNR was calculated to examine the perceived measurement of the cover images and watermark material following watermark embedding.

Table 1. Analysis of reversibility, imperceptibility, and robustness features with no attack

Image Name	Watermarked Image			Logo		Cover Image	
	MSE	PSNR	Payload	NC	BER	NC	BER
Lena	1.4282	46.6168	4096	0.9957	0.0752	0.9953	0.0035
Peppers	2.0797	44.9148	4096	0.9948	0.0827	0.9933	0.0130
Clown	2.8657	43.5925	4096	0.9988	0.0271	0.9997	0.0002
Elain	2.1146	44.9825	4096	0.9994	0.0243	0.9999	0.0001
Goldhill	4.5860	41.5504	4096	0.9990	0.0266	0.9996	0.0003
Pirate	4.4194	41.7112	4096	0.9986	0.0284	0.9995	0.0004
Zelda	0.7552	49.3842	4096	0.9967	0.0623	0.9960	0.0030
Medical Image	4.9077	41.2560	4096	0.9978	0.0344	0.9962	0.0040
Military Image	4.4233	41.0910	4096	0.9965	0.0630	0.9985	0.0015

Table 2. Imperceptibility and Robustness under Distinctive Attacks

Attack	Watermarked Pepper Image		Logo	
	MSE	PSNR	NC	BER
Enhancement technique attacks				
(i)Gaussian Lowpass Filter(3,3)	3.5315	42.6852	0.9648	0.0384
(ii)Median Filter(3,3)	7.8414	39.2209	0.7294	0.3135
(iii)Average Filter(3,3)	10.7780	37.8394	0.7416	0.2983
(iv)Image Sharpening	55.3000	30.7375	0.9546	0.0463
(v)Histogram Equalisation	2.2588	44.6260	0.8946	0.1165
(vi)Gamma Correction(gamma=0.5)	2734.6	13.7958	0.9354	0.0681
Noise addition attack				
(i) Salt & Pepper noise(density=0.5)	110.9691	27.7128	0.8753	0.1272
Geometric transformation attacks				
(i) Rotation(clockwise 2°)	99.941	29.3726	0.8451	0.1763
(ii) Cropping (128x128 by White)	2.0201	45.1110	0.9317	0.0845
(iii) Scaling(zoomout=2,zoomin=0.5)	3.2539	43.0408	0.9174	0.0996
(iv) Cut(20 rows in both up and down)	22.0521	34.7303	0.9249	0.0891
Compression attack				
(i) JPEG Compression (Q=75)	4.0474	42.0930	0.9933	0.0064

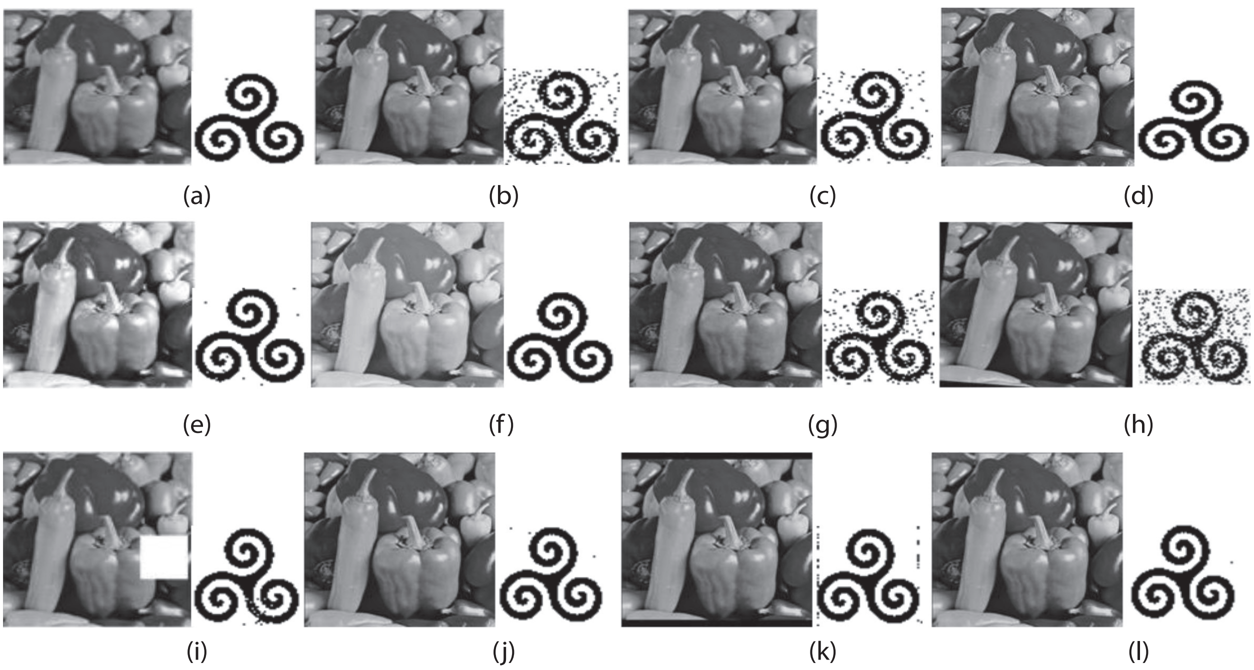


Fig. 5. Attacked Cover Object and Extracted Watermark Images using (**a**)Gaussian Enhancement technique attack (**b**)Median Filter Enhancement technique attack (**c**) Average Filter Enhancement technique attack (**d**) Image Sharpening Enhancement technique attack (**e**)Histogram Equalization Enhancement technique attack (**f**) Gamma correction Enhancement technique attack (**g**) Salt & Pepper Noise addition attack (**h**) Rotation Geometric transformation attack (**i**) Cropping Geometric transformation attack (**j**) Scaling Geometric transformation attack (**k**) Cut Geometric transformation attack (**l**) JPEG Compression attack

Table 1 presents an overview of the experimental outcomes for the suggested watermarking strategy, considering the MSE and PSNR values without affecting the watermarked image in any way. Fig. 4 displays all the watermarked images side by side with the retrieved watermark symbols from the associated watermarked imagery.

(iii) Robustness Measurement: Given its robustness and resilience, the watermarked product should withstand both purposeful and inadvertent attacks aimed at removing the watermark. The robustness of the suggested approach is examined using the normalized cross-correlation (NC) quantity and the bit error rate (BER) between the recovered distorted watermark (without attack) and the original watermark (after using various types of attack). Without any alterations or attacks, the NC and BER measurements of the exerted watermark and the initial watermark are evaluated. In Table 1, these values are displayed.

The watermarked image can be altered illegally while it is being transmitted through the Internet. To provide fair benchmarking and performance assessment, the suggested method is tested against many attackers. Fig. 5 shows each of the watermarked images utilizing different attacks side by side with the recovered logos from the matching watermarked images. Only the experimental results of the 64 × 64 binary logo and one cover image (Pepper image) are shown throughout this study as an example. Table 2 summarizes every attack that is executed using the proposed approach.

(iv) Embedding capacity Measurement: The capacity of a watermark, also known as the watermark payload, is the amount of data implanted as a watermark that can be efficiently extracted on the receiver side without compromising the original data's imperceptibility. The imperceptibility of the watermarked material may be impacted by improving the watermarking scheme's robustness by raising the embedded watermarking payload capacity. This proposed approach modifies a single pair of mid-significant AC components to incorporate a watermark bit within each 8 x 8 deconstructed and non-overlapping block of the DCT-converted host objects.

The imperceptibility of the watermarked object may be impacted by strengthening the robustness of the watermarking method. Thus, to create an effective watermarking method, all three properties must be negotiated. The decomposition block size of the host image can be changed to modify the embedding capability of the suggested method. Table 3 shows the maximum watermark capacity for this approach together with different cover image characteristics.

Table 3. Embedding Capacity of this reversible scheme

Cover Image Size	Host Image Block Size after Decomposition	Watermark Size in Bits
1024 × 1024	8 × 8	2^{14}
1024 × 1024	4 × 4	2^{16}
512 × 512	8 × 8	2^{12}
512 × 512	4 × 4	2^{14}

Comparative Analysis

The suggested approach is contrasted with two current reversible watermarking systems to assess it methodically. In terms of embedded payload (in bits) and PSNR (imperceptibility), this comparative comparison is conducted. Table 4 shows how well the suggested design performs in comparison to the other two systems.

Table 4. Relative study of imperceptibility (PSNR) and embedding capacity (in Bits) of the proposed scheme with some existing techniques

Zhang[18] Lena		Zhang[17] Lena		Proposed Scheme Lena	
Imperceptibility	Embedding Capacity	Imperceptibility	Embedding Capacity	Imperceptibility	Embedding Capacity
39.10	4.4 × 103	37.90	256	46.62	4096

5. CONCLUSIONS

There are several critical challenges associated with current reversible watermarking techniques: (i) most methods are inherently fragile; (ii) the embedding capacity of these algorithms is typically dependent on the signal; (iii) robustness remains a major concern for reversible watermarking techniques; and (iv) the lack of a standardized benchmarking tool for evaluating these schemes. This study proposes a reversible watermarking approach based on histogram modification and inter-block discrete cosine transform (DCT). The proposed system not only ensures reversibility but also demonstrates significant robustness against common image manipulation attacks. Unlike signal-dependent methods, its embedding capacity depends on the size of the DCT block. Experimental results validate the superior performance of the proposed method compared to several existing techniques. The suggested method undergoes a thorough comparison with two existing reversible watermarking systems. Experimental results indicate that it significantly outperforms these current approaches.

6. REFERENCES

[1] S. Roy, A. K. Pal, "A robust blind hybrid image watermarking scheme in RDWT-DCT domain using Arnold scrambling", Multimedia Tools and Applications, Vol. 76, No. 3, 2017, pp. 3577-3616.

[2] F. Mintzer, G. W. Braudaway, "If one watermark is good, are more better?", Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, Vol. 4, Phoenix, AZ, USA, 15-19 March 2019, pp. 2067-2069.

[3] B. Yang, Z. M. Lu, S. H. Sun, "Reversible watermarking in the VQ-compressed domain", Proceedings of the Fifth IASTED International Conference on Visualization, Imaging, and Image Processing, Benidorm, Spain, September 2005, pp. 298-303.

[4] Z. M. Lu, J. X. Wang, B. B. Liu, "An improved lossless data hiding scheme based on image VQ-index residual value coding", Journal of Systems and Software, Vol. 82, No. 6, 2009, pp. 1016-1024.

[5] J. X. Wang, Z. M. Lu, "A path optional lossless data hiding scheme based on VQ joint neighboring coding", Information Sciences, Vol. 179, No. 19, 2009, pp. 3332-3348.

[6] C. C. Chang, C. Y. Lin, Y. H. Fan, "Lossless data hiding for color images based on block truncation coding", Pattern Recognition, Vol. 41, No. 7, 2008, pp. 2347-2357.

[7] B. G. Mobasseri, D. Cinalli, "Lossless watermarking of compressed media using reversibly decodable packets", Signal Processing, Vol. 86, No. 5, 2006, pp. 951-961.

[8] J. Fridrich, M. Goljan, R. Du, "Lossless data embedding—new paradigm in digital watermarking", EURASIP Journal on Applied Signal Processing, Vol. 2002, No. 2, 2002, pp. 185-196.

[9] M. U. Celik, G. Sharma, A. M. Tekalp, E. S. Saber, "Localized lossless authentication watermark (LAW)", Proceedings Volume 5020 of Security and Watermarking of Multimedia Contents V, 2003, pp. 689-698.

[10] U. Celik, G. Sharma, A. M. Tekalp, E. Saber, "Lossless generalized-LSB data embedding", IEEE Transactions on Image Processing, Vol. 14, No. 2, 2005, pp. 253-266.

[11] Y. M. Cheung, H. T. Wu, "A sequential quantization strategy for data embedding and integrity verification", IEEE Transactions on Circuits and Systems for Video Technology, Vol. 17, No. 8, 2007, pp. 1007-1016.

[12] J. D. Lee, Y. H. Chiou, J. M. Guo, "Reversible data hiding based on histogram modification of SMVQ indices", IEEE Transactions on Information Forensics and Security, Vol. 5, No. 4, 2010, pp. 638-648.

[13] L. T. Ko, J. E. Chen, Y. S. Shieh, H. C. Hsin, T. Y. Sung, "Nested quantization index modulation for reversible watermarking and its application to healthcare information management systems", Computational and Mathematical Methods in Medicine, Volume 2012, No. 1, 2012, pp. 839161.

[14] B. Yang, M. Schmucker, W. Funk, C. Busch, S. Sun, "Integer DCT-based reversible watermarking for images using companding technique", Proceedings of the Security, Steganography, And Watermarking of Multimedia Contents VI, Vol. 5306, June 2004, pp. 405-415.

[15] B. Yang, M. Schmucker, M. Niu, C. Busch, S. Sun, "Integer-DCT-based reversible image watermarking by adaptive coefficient modification", Proceedings of the Security, steganography, and watermarking of multimedia contents VII, Vol. 5681, March 2005, pp. 218-229.

[16] G. Xuan, Q. Yao, C. Yang, J. Gao, P. Chai, Y. Q. Shi, Z. Ni, "Lossless data hiding using histogram shifting method based on integer wavelets", International Workshop on Digital Watermarking, Jeju Island, Korea, 8-10 November 2006, pp. 323-332.

[17] X. Zhang, "Reversible data hiding in encrypted image", IEEE Signal Processing Letters, Vol. 1, No. 4, 2011, pp. 255-258.

[18] X. Zhang, "Separable reversible data hiding in encrypted image", IEEE Transactions on Information Forensics and Security, Vol. 7, No. 2, 2011, pp. 826-832.

[19] J. Tian, "Reversible data embedding using a difference expansion", IEEE Transactions on Circuits and Systems for Video Technology, Vol. 13, No. 8, 2003, pp. 890-896.

[20] A. M. Alattar, "Reversible watermark using the difference expansion of a generalized integer transform", IEEE Transactions on Image Processing, Vol. 13, No. 8, 2004, pp. 1147-1156.

[21] J. Y. Hsiao, K. F. Chan, J. M. Chang, "Block-based reversible data embedding", Signal Processing, Vol. 89, No. 4, 2009, pp. 556-569.

[22] C. C. Chang, T. C. Lu, "A difference expansion oriented data hiding scheme for restoring the original host images", Journal of Systems and Software, Vol. 79, No. 12, 2006, pp. 1754-1766.

[23] C. C. Lee, H. C. Wu, C. S. Tsai, Y. P. Chu, "Adaptive lossless steganographic scheme with centralized difference expansion", Pattern Recognition, Vol. 41, No. 6, 2008, pp. 2097-2106.

[24] C. C. Lin, N. L. Hsueh, "A lossless data hiding scheme based on three-pixel block differences", Pattern Recognition, Vol. 41, No. 4, 2008, pp. 1415-1425.

[25] S. Das, A. K. Sunaniya, R. Maity, N. P. Maity, "Efficient FPGA implementation and verification of difference expansion based reversible watermarking with improved time and resource utilization", Microprocessors and Microsystems, Vol. 83, 2021, p. 103732.

[26] H. J. Kim, V. Sachnev, Y. Q. Shi, J. Nam, H. G. Choo, "A novel difference expansion transform for reversible data embedding", IEEE Transactions on Information Forensics and Security, Vol. 3, No. 3, 2008, pp. 456-465.

[27] S. Das, R. Maity, N. P. Maity, "VLSI-based pipeline architecture for reversible image watermarking by difference expansion with high-level synthesis approach", Circuits, Systems, and Signal processing, Vol. 37, 2018, pp. 1575-1593.

[28] Y. Hu, H. K. Lee, J. Li, "DE-based reversible data hiding with improved overflow location map", IEEE Transactions on Circuits and Systems for Video Technology, Vol. 19, No. 2,2008, pp. 250-260.

[29] O. M. Al-Qershi, B. E. Khoo, "High capacity data hiding schemes for medical images based on difference expansion", Journal of Systems and Software, Vol. 84, No. 1, 2011, pp. 105-112.

[30] F. H. Hsu, M. H. Wu, S. J. Wang, C. L. Huang, "Reversibility of image with balanced fidelity and capacity upon pixels differencing expansion", The Journal of Supercomputing, Vol. 66, No. 2, 2013, pp. 812-828.

[31] K. Jawad, A. Khan, "Genetic algorithm and difference expansion based reversible watermarking for relational databases", Journal of Systems and Software, Vol. 86, No. 11, 2013, pp. 2742-2753.

[32] S. Weng, Y. Zhao, J. S. Pan, R. Ni, "Reversible data hiding using the companding technique and improved DE method", Circuits, Systems, and Signal Processing, Vol. 27, No. 2, 2008, pp. 229-245.

[33] S. Das, A. K. Sunaniya, R. Maity, N. P. Maity, "Parallel hardware implementation of efficient embedding bit rate control based contrast mapping algorithm for reversible invisible watermarking", IEEE Access, Vol. 8, 2020, pp. 69072-69095.

[34] W. Hong, J. Chen, T. S. Chen, "Blockwise reversible data hiding by contrast mapping", Information Technology Journal, Vol. 8, No. 8, 2009, pp. 1287-1291.

[35] D. M. Thodi, J. J. Rodríguez, "Expansion embedding techniques for reversible watermarking", IEEE Transactions on Image Processing, Vol. 16, No. 3, 2007, pp.721-730.

[36] P. Tsai, Y. C. Hu, H. L. Yeh, "Reversible image hiding scheme using predictive coding and histogram shifting", Signal Processing, Vol. 89, No. 6, 2009, pp.1129-1143.

[37] H. W. Tseng, C. P. Hsieh, "Prediction-based reversible data hiding", Information Sciences, Vol. 179, No. 14, 2009, pp. 2460-2469.

[38] V. Sachnev, H. J. Kim, J. Nam, S. Suresh, Y. Q. Shi, "Reversible watermarking algorithm using sorting and prediction", IEEE Transactions on Circuits and Systems for Video Technology, Vol. 19, No. 7, 2009, pp. 989-999.

[39] D. Coltuc," Improved embedding for prediction-based reversible watermarking", IEEE Transactions on Information Forensics and Security, Vol. 6, No. 3, 2011, pp. 873-882.

[40] D. Coltuc, "Low distortion transform for reversible watermarking", IEEE Transactions on Image Processing, Vol. 21, No. 1, 2011, pp. 412-417.

[41] G. Feng, L. Fan, "Reversible data hiding of high payload using local edge sensing prediction", Journal of Systems and Software, Vol. 85, No. 2, 2012, pp. 392-399.

[42] G. Feng, Z. Qian, N. Dai, "Reversible watermarking via extreme learning machine prediction", Neurocomputing, Vol. 82, 2012, pp. 62-68.

[43] H. T. Wu, J. Huang, "Reversible image watermarking on prediction errors by efficient histogram modification", Signal Processing, Vol. 92, No. 12, 2012, pp. 3000-3009.

[44] X. Chen, X. Sun, H. Sun, Z. Zhou, J. Zhang, "Reversible watermarking method based on asymmetric-histogram shifting of prediction errors", Journal of Systems and Software, Vol. 86, No. 10, 2013, pp. 2620-2626.

[45] H. Y. Leung, L. M. Cheng, F. Liu, Q. K. Fu, "Adaptive reversible data hiding based on block median preservation and modification of prediction errors", Journal of Systems and Software, Vol. 86, No. 8, 2013, pp. 2204-2219.

[46] X. Li, B. Yang, T. Zeng, "Efficient reversible watermarking based on adaptive prediction-error expansion and pixel selection", IEEE Transactions on Image Processing, Vol. 20, No. 12, 2011, pp. 3524-3533.

[47] X. Li, J. Li, B. Li, B. Yang, "High-fidelity reversible data hiding scheme based on pixel-value-ordering and prediction-error expansion", Signal Processing, Vol. 93, No. 1, 2013, pp.198-205.

[48] B. Ou, Y. Zhao, R. Ni, "Reversible watermarking using optional prediction error histogram modification", Neurocomputing, Vol. 93, 2012, pp. 67-76.

[49] B. Ou, X. Li, Y. Zhao, R. Ni, "Reversible data hiding based on PDE predictor", Journal of Systems and Software, Vol. 86, No. 10, 2013, pp. 2700-2709.

[50] X. Shi, D. Xiao, "A reversible watermarking authentication scheme for wireless sensor networks", Information Sciences, Vol. 240, 2013, pp. 173-183.

[51] X. Zhang, "Reversible data hiding with optimal value transfer." IEEE Transactions on Multimedia, Vol. 15, No. 2, 2012, pp. 316-325.

[52] L. Luo, Z. Chen, M. Chen, X. Zeng, Z. Xiong, "Reversible image watermarking using interpolation technique", IEEE Transactions on Information Forensics and Security, Vol. 5, No. 1, 2009, pp. 187-193.

[53] W. Hong, T. S. Chen, "Reversible data embedding for high quality images using interpolation and reference pixel distribution mechanism", Journal of Visual Communication and Image Representation, Vol. 22, No. 2, 2011, pp. 131-140.

[54] C. De Vleeschouwer, J. F. Delaigle, B. Macq, "Circular interpretation of bijective transformations in lossless watermarking for media asset management", IEEE Transactions on Multimedia, Vol. 5, No. 1, 2003, pp. 97-105.

[55] Z. Ni, Y. Q. Shi, N. Ansari, W. Su, "Reversible data hiding", IEEE Transactions on Circuits and Systems for Video Technology, Vol. 16, No. 3, 2006, pp. 354-362.

[56] Z. Ni, Y.Q. Shi, N. Ansari, W. Su, Q. Sun, X. Lin, "Robust lossless image data hiding designed for semi-fragile image authentication", IEEE Transactions on Circuits and Systems for Video Technology, Vol. 18, No. 4, 2008, pp. 497-509.

[57] H. W. Tseng, C. C. Chang, "An extended difference expansion algorithm for reversible watermarking", Image and Vision Computing, Vol. 26, No. 8, 2008, pp. 1148-1153.

[58] C. C. Lin, W. L. Tai, C. C. Chang, "Multilevel reversible data hiding based on histogram modification of difference images", Pattern Recognition, Vol. 41, No. 12, 2008, pp. 3582-3591.

[59] Y. C. Li, C. M. Yeh, C. C. Chang, "Data hiding based on the similarity between neighboring pixels with reversibility", Digital Signal Processing, Vol. 20, No. 4, 2010, pp. 1116-1128.

[60] K. S. Kim, M. J. Lee, H. Y. Lee, H. K. Lee, "Reversible data hiding exploiting spatial correlation between sub-sampled images", Pattern Recognition, Vol. 42, No. 11, 2009, pp. 3083-3096.

[61] X. Gao, L. An, X. Li, D. Tao, "Reversibility improved lossless data hiding", Signal Processing, Vol. 89, No. 10, 2009, pp. 2053-2065.

[62] W. L. Tai, C. M. Yeh, C. C. Chang, "Reversible data hiding based on histogram modification of pixel differences", IEEE Transactions and Circuits and Systems for Video Technology, Vol. 19, No. 6,2009, pp. 906-910.

[63] X. Li, B. Li, B., B. Yang, T. Zeng, “General framework to histogram-shifting-based reversible data hiding", IEEE Transactions on Image Processing, Vol. 22, No. 6, 2013, pp. 2181-2191.

[64] X. Li, W. Zhang, X. Gui, B. Yang, "A novel reversible data hiding scheme based on two-dimensional difference-histogram modification", IEEE Transac-

tions on Information Forensics and Security, Vol. 8, No. 7, 2013, pp. 1091-1100.

[65] Q. Pei, X. Wang, Y. Li, H. Li, "Adaptive reversible watermarking with improved embedding capacity", Journal of Systems and Software, Vol. 86, No. 11, 2013, pp. 2841-2848.

[66] A. Khan, S. A. Malik, "A high capacity reversible watermarking approach for authenticating images: Exploiting down-sampling, histogram processing, and block selection", Information Sciences, Vol. 256, 2014, pp. 162-183.

[67] M. Kutter, F. A. Petitcolas, "Fair benchmark for image watermarking systems", Proceedings of Electronic Imaging'99, Security and Watermarking of Multimedia Contents, Vol. 3057, San Jose, CA, USA, 25-27 January 1999, pp. 226-239.

[68] S. Voloshynovskiy, S. Pereira, T. Pun, J. J. Eggers, J. K. Su, "Attacks on digital watermarks: classification, estimation based attacks, and benchmarks", IEEE Communications Magazine, Vol. 39 No. 8, 2001, pp. 118-126.

Customer In-Store Behavior Analysis Using Beacon Data at a Home Improvement Retailer

Original Scientific Paper

Ayla Gülcü*
Bahçeşehir University,
Faculty of Engineering and Natural Sciences,
Department of Software Engineering
Istanbul, TÜRKIYE
ayla.gulcu@bau.edu.tr

İnanç ONUR
Patika Global Technology,
Istanbul, TÜRKIYE
inanc.onur@patikaglobal.com

Sümeyra Öztop
Patika Global Technology,
Istanbul, TÜRKIYE
sumeyra.oztop@patikaglobal.com

*Corresponding author

Enes Uğurlu
Patika Global Technology,
Istanbul, TÜRKIYE
enes.ugurlu@patikaglobal.com

Remzi Emre Sain
Patika Global Technology,
Istanbul, TÜRKIYE
remzi.sain@patikaglobal.com

Abstract *– In this study, we aimed to analyze the in-store behavior of customers at a home improvement retail company using data collected from Bluetooth Low Energy beacon devices installed on shelves and shopping carts within a selected store. The beacons were strategically placed on store shelves to ensure complete coverage, leaving no blind spots. To cover 18 departments spanning a total area of approximately 4,800 square meters, 99 beacons were deployed. The duration of stay in each department, the order of visits, and the absolute visit date and time were recorded in the database. To investigate the relationship between in-store behavior and purchase data, we combined customers' behavioral data with their purchase information. Correlation analysis revealed a positive relationship between visit duration and purchase amount, particularly in the Floor Deco, Paint, and Taps departments. Additionally, we visualized store-wide data using a network diagram, highlighting key shopping areas, customer flow patterns, and high-revenue departments. The problem was also formulated as a multi-class classification task, and LSTM and XGBoost algorithms were applied for comparative analysis. Experiments were conducted on both the original dataset and a cleaned version, utilizing two distinct data modeling approaches: one based solely on sequential department visits and another incorporating visit duration. The results showed that both models performed similarly on the noisy dataset, indicating that adding duration information did not improve learning. However, when trained on the cleaned dataset where short-duration visits were removed, LSTM models outperformed XGBoost, demonstrating a stronger ability to capture meaningful sequential patterns. These findings highlight the potential of BLE beacon technology in retail analytics, offering deeper insights into customer behavior and informing data-driven decision-making for store optimization and personalized marketing. Future work will focus on expanding the dataset and refining predictive models to further enhance the accuracy and applicability of in-store behavior analysis.*

Keywords: *BLE Beacon localization, data analysis, internet of things, machine learning*

Received: December 28, 2024; Received in revised form: February 21, 2025; Accepted: March 31, 2025

1. INTRODUCTION

Companies can gain valuable insights by analyzing customer behavior data. By leveraging data mining and machine learning techniques, businesses can identify high-value customers, capture their personalized needs, and implement effective marketing strategies.

Zhao [1] highlights that personalized marketing, based on extensive customer data analysis, can significantly enhance marketing campaigns, with conversion rates being 40% higher compared to non-personalized approaches. Traditionally, customers' purchase histories and Point of Sales (POS) data have been the primary tools for analyzing customer segments, especially when combined with demographic or geographic data. However, incorporating behavioral factors such as shopping duration, wandering routes, and visiting patterns can provide deeper insights into customer preferences and behaviors.

Behavioral data can help answer critical business questions, as discussed in [2], such as: What are the most frequent paths taken by customers? Which areas are skipped? Where do they spend the majority of their time? Do they primarily navigate the outer ring of the store, or do they focus on specific sections?

The Internet of Things (IoT) has become an essential tool, particularly in the retail sector, for gathering customer behavior data (see [3] for a review of the digital transformation across sectors). IoT sensors placed on store shelves or shopping carts can be utilized to analyze customers' shopping behaviors. Additionally, smart mobile applications can assist customers with in-store navigation. Various IoT technologies are employed for in-store customer tracking and localization, including: i) Bluetooth Low Energy (BLE) Beacon technology, ii) Ultra-Wideband (UWB) technology, iii) Wayfinding technologies, and iv) Camera and image processing technologies. BLE Beacon technology is particularly favored for its low energy consumption and cost-effectiveness.

Ke et al. [4] compare various Radio-Frequency (RF)-based indoor positioning technologies. They propose a BLE Beacon-based localization system designed to detect human locations in smart homes for power management purposes. Another application of indoor positioning is discussed in [5]. Spachos and Plataniotis [6] utilize BLE Beacons in an IoT-based smart museum to deliver interactive and personalized museum tours. Similarly, Shipkovenski et al. [7] present an indoor positioning system for hospital environments, enabling patient tracking and preventing unauthorized departures. Pangriya [8] provides a comprehensive literature review on Beacon technology, offering valuable insights into its future in the retail sector. The author also presents the results of a SWOT analysis (Strengths, Weaknesses, Opportunities, and Threats), conducted through two-phase interviews with experts from the technology and retail industries. Among the strengths of Beacon technology are its cost-effectiveness and ability to function indoors with high precision. However, its primary weakness is the requirement for customers to keep Bluetooth enabled and the store's mobile application running while in the store.

In this study, we analyze the in-store behavior of customers at a home improvement retailer using data collected from BLE beacon devices installed on shelves and shopping carts. We integrate this behavior data with purchase data to uncover key customer characteristics. We formulate the problem as a multi-class classification task and apply two distinct algorithms comparatively using two different models. Our approach differs from prior studies by introducing a novel data modeling technique that integrates both sequential department visit data and visit duration. Unlike studies that focus on a single department, our analysis spans 18 departments, providing a more comprehensive view of customer behavior. Unlike studies that focus on customer behavior within a single department, this research considers data spanning 18 departments, providing a more comprehensive understanding of customer behavior.

The remainder of the paper is organized as follows: Section 2 reviews related studies. Section 3 details beacon installation, data collection, and modeling approaches. Section 4 presents correlation analysis, visualizations, and algorithm comparisons. Finally, Section 5 concludes with key findings and remarks.

2. LITERATURE REVIEW

This section focuses on studies that utilize data collected from Beacon devices to infer user or customer behavior. While earlier studies primarily relied on Radio Frequency Identification (RFID) tags, recent advancements have led to the adoption of Bluetooth Low Energy (BLE) Beacons due to their cost-effectiveness, ease of deployment, and ability to function efficiently in real-world retail environments.

One of the earliest studies on customer in-store analysis using RFID was conducted by Larson et al. [2], who identified frequent shopping routes in a grocery store. They analyzed 27,000 shopper paths, with path lengths varying from short 2-minute routes to longer 2-hour shopping patterns. Their findings demonstrated that certain store layouts influence customer movement, providing insights into strategic shelf placements.

Recent advancements have led to the adoption of BLE Beacons, which offer a more scalable and flexible approach to in-store tracking. Unlike RFID, BLE Beacons can continuously monitor customer movement without requiring direct scanning or RFID-tagged objects. Instead, they rely on smartphone interactions or strategically placed beacons to track customer flow. Jain et al. [9] propose a comprehensive framework for evaluating the adoption and readiness of beacon technology. They highlight a significant gap in the field of proximity marketing and beacon technology and suggest that their study can help retail managers better understand the added value of implementing this technology.

Lemsieh and Abarar [10] conducted a study on customer perceptions of proximity marketing using BLE beacon technology in Moroccan malls. The results indicate that customers are eager to adopt this technology, as it helps them locate stores independently, provided their privacy concerns are adequately addressed. Due to these concerns, most studies focus on using BLE beacon technology for indoor localization of customers or personnel within a facility rather than for delivering personalized experiences. Garcia and Inoue [11] presents a study on the indoor localization of caregiving and nursing staff in a nursing Care facility using BLE Beacon technology. Spachos and Plataniotis [6] demonstrated the effectiveness of BLE Beacons in guiding museum visitors through location-aware tours. In another study, Shende et al. [12] utilized BLE Beacons to analyze customer movements in a shopping mall, enabling real-time personalized promotions based on movement patterns.

For personalized marketing, BLE beacon technology which can be seamlessly integrated with mobile phones is essential. In this study, we used this technology to collect customer in-store behavior data in a home improvement retailer. Analyzing customer behavior data helped to answer many business questions. We also combined the behavior data with the purchase data in order to investigate the relationship between them. We formulated the problem as a multi-class classification task, and employed LSTM and XGBoost algorithms. We introduced a novel data modeling approach that captures both the sequence of department visits and the duration of those visits.

Most studies in the literature that apply machine learning algorithms to analyze customer behavior data rely on data collected using RFID technology. More importantly, many of these studies focus on a single department within a store, whereas our study utilizes data from 18 departments. For instance, Zhao et al. [13] proposed a sequential classification-based model that integrates RFID data with point-of-sale (POS) data to classify and identify consumers' purchasing behavior, but their analysis was limited to a specific island area of a supermarket. Similarly, Zuo et al. [14] combined RFID-based customer visit data with POS transactions in a Japanese supermarket, focusing solely on the fish section. Extending this approach, Zuo et al. [15] placed RFID tags on shopping carts and installed receptors beneath store shelves to track customers' real-time interactions with different product aisles, though their study was confined to the bread section.

There are also studies that employ classification algorithms to analyze customer behavior in the e-commerce domain. For example, Li [16] utilized customers' past product browsing activity data to construct a classification model based on the XGBoost algorithm to predict the customer purchasing behavior in the e-commerce domain. Similarly, Park et al. [17] applied the same algorithm to predict the product a customer would purchase based on their browsing activity.

Our study differs from existing literature by incorporating sequential data into the classification task. Additionally, we introduce a novel modeling approach that integrates both sequential data and numerical duration information. Furthermore, rather than focusing on a single department, we utilize data from 18 departments, enabling a comprehensive analysis across the entire dataset.

3. METHODOLOGY

In this section, we first describe the BLE beacon installation process within the retail store. Next, we outline the localization and data collection procedures. Finally, we detail the data modeling approach employed and the machine learning algorithms applied in this study.

3.1. BEACON INSTALLATION AND DATA COLLECTION

Noise Filtering Process: Beacons transmit signals at regular intervals using Bluetooth Low Energy (BLE) technology. In this study, the system includes two types of beacons: shelf beacons and cart beacons. Shelf beacons act as transmitters, emitting signals within a defined area and sending them to nearby devices. Cart beacons function as receivers, capturing and recording signals emitted by the shelf beacons that correspond to the aisle where the customer's cart is located. This setup enables the system to accurately track the customer's location within the store.

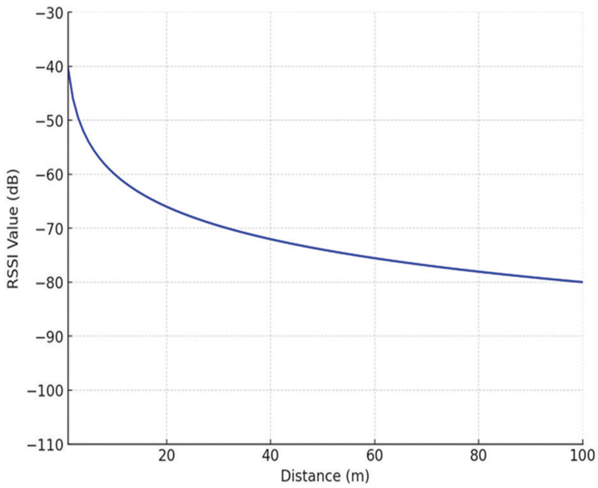


Fig. 1. The relationship between RSSI values and distance

The Received Signal Strength Indicator (RSSI) value, measured in decibels (dB), represents signal strength and is used to estimate the distance between a device and a beacon. As the distance between the transmitting beacon and the receiving device increases, the RSSI value decreases (becomes more negative) due to signal attenuation, a phenomenon where the signal's energy dissipates over distance. Fig. 1 illustrates how RSSI values vary with distance. RSSI data obtained from beacon devices provides a measure of signal strength; however, it is susceptible to environmental factors. This phenomenon, known as noise, can cause fluctuations in signal strength, reducing accuracy. Noise may originate from various sources, such as other electronic devices, structural obstacles (e.g., walls and shelves), human crowds, and even continuous variations in the distance between the beacon and the receiver.

The optimal placement of transmitter beacon devices on store shelves, which communicate with receiver beacon devices in customers' shopping carts, is crucial for ensuring effective communication. Extensive experiments were conducted in a controlled physical environment to determine the optimal signal range while minimizing interference from environmental noise. Table 1 summarizes the data collected by receiver beacons during the experimental visits.

Table 1. Experiments to determine the RSSI filter value

Id	RSSI Filter (dB)	Beacon Region	Interaction Duration	Diff Time
D1	-84	Lighting	1727	899
D1	-84	Promo	1727	899
D2	-82	Lighting	1160	922
D2	-82	Promo	1160	922
D3	-80	Lighting	402	437
D3	-80	Promo	402	437

Two departments, namely Lighting and Promo, were visited three times during the experiments. D1, D2, and D3 represent these visits, each conducted under different RSSI filter values. The Interaction Duration column indicates the total time the beacon received signals from transmitter beacons in the surrounding area, while the Diff Time column reflects the duration the beacon received signals irrespective of the number of beacons nearby. In the experimental visit D1 with a filter value of –84 dB, the Interaction Duration value was significantly higher than the Diff Time value. This discrepancy suggests that the beacon was receiving noisy signals from beacons installed in other departments. In contrast, experimental visit D3 with a filter value of –80 dB yielded similar values for Interaction Duration and Diff Time, indicating minimal noise interference. Thus, a filter value of –80 dB was determined to be optimal for this environment.

Beacon Localization: After determining a noise filter value of –80 dB, the distance between any two beacons needed to be defined. Since the areas of the departments varied, the number of beacons installed in each department also differed. For a given department, beacon placement was performed manually by measuring signal strength as follows: The tangent point was identified as the location where the RSSI value, measured while moving away from the first beacon, dropped to the predefined threshold of –80 dB. From this tangent point, a second beacon was gradually moved forward, with its signal strength at the tangent point continuously measured. Once the measured value dropped below the threshold, the beacon was fixed at its current position. Subsequent devices were positioned similarly, ensuring that their coverage areas were tangent to one another. Care was taken to ensure that blind spots - areas not covered by any beacon- did not exceed 2 meters. Depending on the department, beacons were placed approximately 8 to 14 meters apart. In total, 99 beacons were installed to provide coverage for an area of approximately 4,800 square meters.

Data Collection: As a shopping cart equipped with a beacon moves around the store, it records the MAC addresses of the beacons it interacts with, along with the interaction start and end times, connection durations, and RSSI values. This data is stored in the database, with the MAC addresses used to identify the departments visited by the cart, thereby mapping the shopping route. Checkout beacons and gateways were employed to match carts with specific checkout counters. Once these cart-checkout matches were established, cart-receipt pairings were conducted to associate the shopping route data with the corresponding purchase details for each cart.

This process allowed the data collected from a customer's cart beacon during their in-store journey to be matched with their shopping details and stored in the database for further processing. At the end of each day, the data was inspected for noise. To clean this noisy data, beacon connection durations and RSSI value information were analyzed. Finally, the departments visited during each trip were identified using the MAC addresses of the shelf beacons. This process is illustrated in Fig. 2.

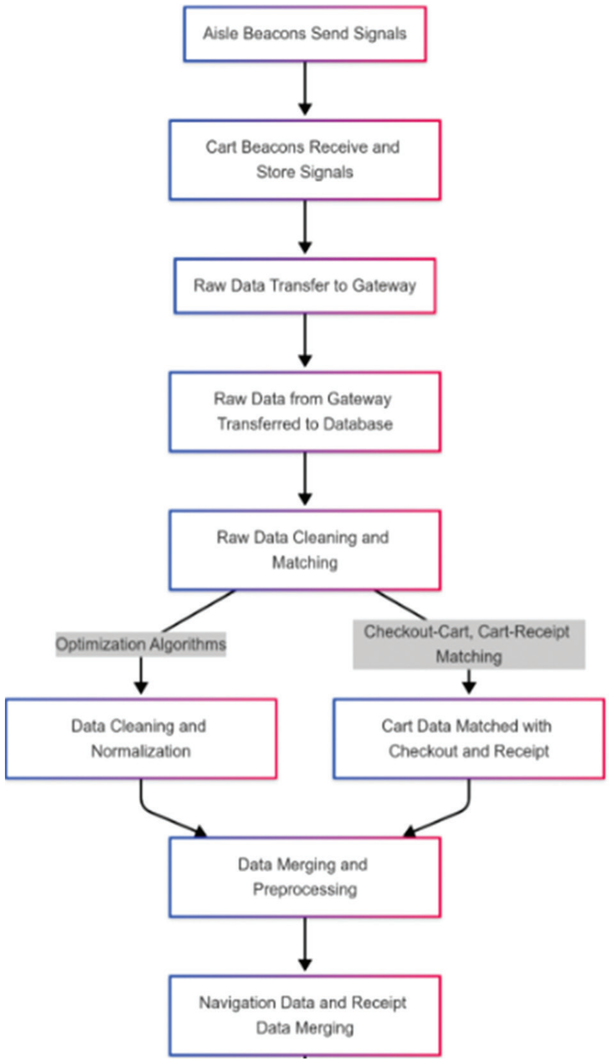


Fig. 2. Illustration of the steps in the data collection and analysis phase

3.2. DATASET

Dataset Format: This study primarily utilizes two types of tables to store customer data. The first table, illustrated in Table 2, records customer visit data. Each customer visit begins with a signal from the entrance beacon and concludes with a signal from the register beacon.

Table 2. Customer behavior Data

User id	Session id	Time	Dep. id	Duration (min.)
U1	S1	2023-07-31 15:00:10	Entrance	0
		2023-07-31 15:00:10	A	9
		2023-07-31 15:09:20	D	21
		2023-07-31 15:30:05	Z	5
		2023-07-31 15:35:05	Register	0
U2	S1	2023-07-31 15:01:10	Entrance	0
		2023-07-31 15:01:10	B	14
		2023-07-31 15:15:20	E	25
		2023-07-31 15:40:30	F	10
		2023-07-31 15:50:30	Register	0
...	...	...	...	...

The data collected from cart beacons undergo preprocessing steps to remove irrelevant information. For instance, if a customer merely passes through a department without spending time there, the beacon signals for that department are excluded. Table 2 provides an example with two visit sessions performed by two different customers. Customer U1 initially visits department D, spends 9 minutes there, and then proceeds to department Z.

The second type of data pertains to customer point-of-sale (POS) information, as shown in Table 3. Each row represents the POS data for a single customer visit, including the products purchased during that session and the total amount spent. To derive meaningful insights from customer behavior, the POS data table is combined with the customer visit data table. All subsequent analyses are conducted using this integrated dataset.

Dataset Details: In this subsection, we present details about the real-world data collected from a selected branch of the home improvement company. The dataset includes customer data recorded between January 9, 2024, and June 10, 2024. Due to a limited number of available shelf beacons, the study focused on 18 selected departments: Lighting, Window Deco, Floor Deco, Wall Deco, Bathroom Accessories, Tiles, Kitchen Accessories, Gardening, Hardware, Heating, Tools, Houseware Storage, Plumbing, Paint, Taps, Furniture, Electrical, and Hard Flooring. Although the entrance and register are not considered departments, they are included in each visit's data. After completing data cleaning and preprocessing phases, the combined dataset, illustrated in Fig. 3, was created. The column labeled checkout receipt id was used to merge the customer visit data with the POS data. For instance, the row with a checkout receipt id value of 316507 indicates that the customer visited six departments. The 'categories' column in the merged dataset table contains department IDs. The 'categories_x' column originates from the customer visit data table and represents the departments visited by a customer, while the 'categories_y' column comes from the POS data and indicates the departments where the customer made purchases.

Table 3. Customer POS Data

User id	Session id	Time	Dep. id	Duration (min.)
U1	S1	2023-07-31 15:35:05	P1, P1, P8, P9, P70	122
U1	S2	2023-07-31 15:50:30	P4, P16, P16	36
U2	S1	2023-08-02 14:15:50	P2	50
...	...	...	...	...

The sequence in which these departments were visited is preserved in the table, along with the duration of the visit (in seconds) for each department. Instead of listing specific product names, the table includes the product categories purchased during the visit. This category information was used to map purchases to corresponding department names. In the example row, the customer visited departments with category IDs 49, 7, 47, 200, 54, and 300 but made purchases only in the department with category ID 49. The purchase amounts associated with each category are also provided, enabling the total purchase amount to be easily computed. During the data collection process, no customer identity-related information was collected; therefore, the data may include multiple visits from the same customer. The final combined dataset comprises 1,447 rows, each representing a distinct customer visit session.

3.3. CLASSIFICATION ALGORITHMS

In this study, we have used two well-known classification approaches comparatively. The first method we employ is a Long Short-Term Memory (LSTM) network, which is a specialized type of Recurrent Neural Networks (RNNs). RNNs are a class of neural networks designed to process sequential data by maintaining a hidden state that captures temporal dependencies. LSTM networks, a specialized type of RNN, address the vanishing gradient problem by incorporating gating mechanisms, namely, input, output, and forget gates to selectively update and retain information over long sequences [18].

While a single-layer LSTM effectively captures sequential dependencies, its ability to extract complex hierarchical features is limited. To enhance performance, researchers often use stacked LSTMs (deep LSTMs), where multiple LSTM layers are stacked, with the hidden state of one layer serving as the input to the next. Another variant, the Bidirectional LSTM (BiLSTM), improves performance by processing sequences in both forward and backward directions. LSTMs are widely used in tasks such as time-series forecasting, natural language processing, and sequence classification due to their ability to model long-term dependencies effectively.

The second method that we employ is Extreme Gradient Boosting (XGBoost), which is a scalable and efficient implementation of gradient-boosted decision trees, designed for both classification and regression tasks [19]. It employs advanced regularization techniques like Lasso regression (L1 regularization) and Ridge regression (L2 regularization) to prevent overfitting and uses an optimized gradient descent approach to minimize loss iteratively. Traditional boosting algorithms, such as Gradient Boosting Machines (GBM), optimize the loss function using first-order gradient information. In contrast, XGBoost improves upon this by incorporating second-order gradient information, which includes both the gradient (first derivative) and the Hessian (second derivative). This allows for better approximation of the loss function. XGBoost is renowned for its speed and accuracy in handling structured data and has been widely adopted in machine learning competitions and real-world applications.

In this study, we formulate our problem as a multi-class classification task, aiming to predict customer sales volumes based on their store visit data. Input vectors representing customer visits were constructed using two distinct modeling approaches, each of which is detailed in this section.

Model 1: The dataset used in this study captures in-store customer behavior along with Point-of-Sale (POS) data. Each row represents a unique shopping session associated with a checkout receipt ID. For instance, the first row (Receipt ID: 321787) indicates that the customer visited 8 departments during their session. The sequence of department visits in this example is as follows: 200, 49, 300, 36, 300, 54, 60, and 54. Each department is represented using an 18-dimensional vector derived from one-hot encoding. Since there are 18 unique departments in total, each department is assigned a unique integer ID ranging from 0 to 17. Examples of department encoding are shown in Table 4.

Each customer session consists of a sequence of department visits. The encoded one-hot vectors for each visited department are concatenated to create the input representation. For example, a session with the sequence: "Kitchen Accessories → Gardening → Tools → Hardware → Tools" would be represented as a matrix of dimensions 5×18. All input vectors were padded to match the length of the longest session. The maximum observed session length is 56 departments. Shorter sessions were padded with the following zero vector:

[0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0]

Thus, each session is represented as a 56×18 matrix. As the dataset contains a total of 1,447 shopping sessions, the complete dataset is represented as a 3-dimensional matrix with dimensions (1447, 56, 18) after encoding and padding.

Table 4. Model 1 Input Representation

Dep. Name	Id	One-Hot Encoding
Kitchen Accessories	0	[1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0]
Heating	1	[0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0]

Model 2: Model 1 preserves the department visit order; however, it does not consider the duration of the visit. In Model 2, the duration of stay in each department is incorporated into the encoding process. A repetition interval of 100 seconds was defined to model department visit durations. The total visit duration for each department was divided by this interval to determine the number of repetitions as shown in Table 5.

Table 5. Model 2 Input Representation

Dep. Id	Visit Dur. (sec.)	Repetition Interval (100s)	Transformed Sequence
54	222	3	[54, 54, 54]
37	36	1	[37]
18	110	2	[18, 18]

checkout_receipt_id	start_time	department_cnt	departments	categories_x	durations	categories_y	unit_price	quantity	sales_amount
321787	2024-06-07 11:26:15	8	Kitchen Acc-Gardening-Tools-Hardware-Tools-Lig...	200-49-300-36-300-54-60-54	8-836-176-100-270-2-18-16	49-300-60-60-60-60-300	949.9-71.99-79.99-37.99-304.9-79.99-31.99	1-1-1-1-1-1-1	949.9-71.99-79.99-37.99-304.9-79.99-31.99
321427	2024-06-10 14:56:07	7	Gardening-Kitchen Acc-Taps-Bathroom&Accessorie...	49-200-33-37-86-402-54	52-882-6-32-132-2-22	49-49	599.9-499.9	1-3	599.9-1499.7
316507	2024-06-09 12:12:53	6	Gardening-Electrical-Houseware&Storage-Kitchen...	49-7-47-200-54-300	932-84-14-48-222-98	49	112.9	1	112.9
308767	2024-01-12 12:27:14	12	Lighting-Wall Deco-Paint-Bathroom&Accessories-...	54-400-60-37-37-200-47-200-200-50-37-54	222-18-56-40-134-132-54-132-28-134-16-192	56-200-54-17,54-60	399.9-249.9-54.99-39.99-24.99	1-1-2-4-1	399.9-249.9-109.98-159.96-24.99
307807	2024-05-08 11:59:31	6	Lighting-Gardening-Houseware&Storage-Bathroom&...	54-49-47-37-54-7	252-78-10-242-82-140	54-47-7-54	149.9-32.99-239.9-149.9	2-1-1-2	299.8-32.99-239.9-299.8

Fig. 3. Combined data table where each row refers to a single customer session behavior and POS data

After transforming the sequences based on durations, the same one-hot encoding and padding steps as in Model 1 were applied. For example, a session with department IDs [54, 37, 18] and durations [222, 36, 110] would be transformed into: [54, 54, 54, 37, 18, 18, 50, 50]. The final input is a padded matrix with dimensions 56×18.

The duration-adjusted dataset contains 1,447 sessions with a maximum transformed session length of 119 departments. The complete input matrix dimensions are: (1447, 119, 18)

This enhanced encoding approach ensures that both visit order and duration are captured effectively, providing a richer representation for sequential modeling tasks.

4. RESULTS AND DISCUSSION

In this section, we first analyze the customer data, including both visit data and POS data, to derive meaningful insights. Subsequently, we present a comparative evaluation of the LSTM and XGBoost algorithms using the two models described in Section 3.

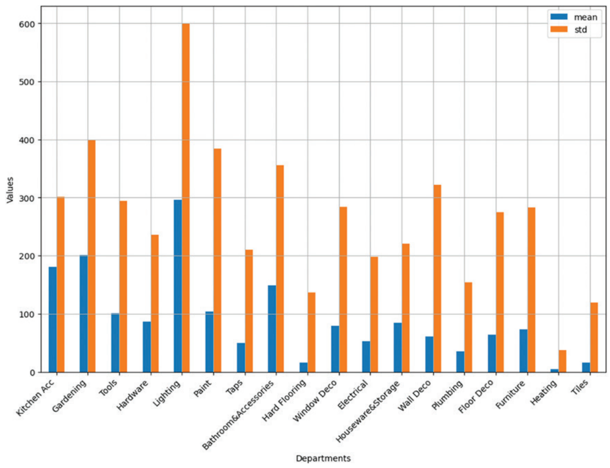


Fig. 4. Department visit duration average and standard deviation

4.1. CORRELATION ANALYSIS

We present the visit duration statistics for each department in Fig. 4. Notably, the 'Lighting' department has the highest average visit duration among all departments. However, it also exhibits the greatest standard deviation, indicating substantial variation in the time customers spend in this section. This suggests that while some customers may spend considerable time in the 'Lighting' department, others may pass through more quickly, resulting in a wide range of visit durations.

The purchase amount statistics for each department are detailed in Fig. 5. The 'Lighting' department has the highest average shopping amount, making it the top-performing section in this regard. Conversely, the 'Taps' department shows a notably high standard deviation, suggesting considerable variability in the prices of purchases made in this section.

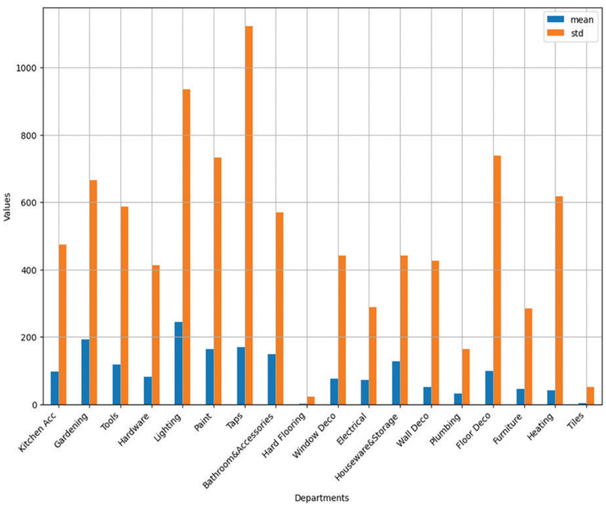


Fig. 5. Department sales amount average and standard deviations

To examine the relationship between the duration of stay in each department and the corresponding purchase amount, we conducted a correlation analysis using Pearson correlation coefficients. A high correlation coefficient indicates a positive relationship between the time spent in a department and the purchase amount. As shown in Fig. 6, the Floor Deco, Paint, and Taps departments exhibit a strong positive correlation, where increased time spent is associated with higher recorded shopping amounts.

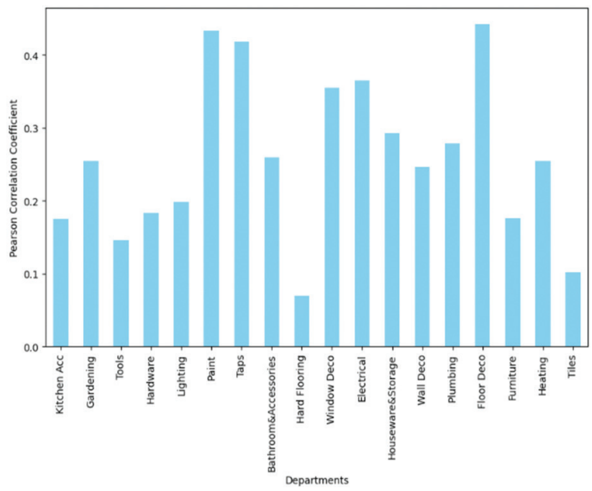


Fig. 6. Correlation between visit duration and purchase amount

4.2. VISUALIZATION

Heatmaps were utilized to visualize the most visited corridors better within the beacon installed areas of the store. The heatmap, based on beacon data collected on June 1, 2024, is shown in Fig. 7. Five different colors were employed on the map: light yellow represents the least dense corridors, while dark red indicates the most dense corridors. We have also created a network graph shown in Fig. 8 using the information from the entire dataset as follows:

- Each node represents a department.
- The size of each node corresponds to the total time spent in that department.
- An edge is drawn between departments with transitions, with edge weights determined by the number of transitions.
- The color of each node corresponds to the total shopping amount recorded in that department.

The network graph indicates that the most frequent transitions occur between the Lighting and Window Deco departments. The Lighting department, represented in the darkest color, records the highest total shopping amount and is also the department where the most time was spent overall. Similarly, the Gardening department shares comparable characteristics with the Lighting department in terms of total time spent and total shopping amount.

4.3. MULTI-CLASS CLASSIFICATION

We approached the problem as a multi-class classification task, where the total purchase amount at the end of each visit session was categorized and used as the response variable. Following the methodology in [13], we carefully analyzed statistical distributions and the histogram of total sales amounts (Fig. 9) to define appropriate category boundaries. Accordingly, three distinct categories for total sales were defined as follows:

- 0 - Low volume: (0, 1000)
- 1 - Medium volume: [1000, 10000)
- 2 - High volume: [10000, 100000)

We applied both algorithms to the input data generated using Model 1, which preserves the order of department visits but does not account for visit duration.

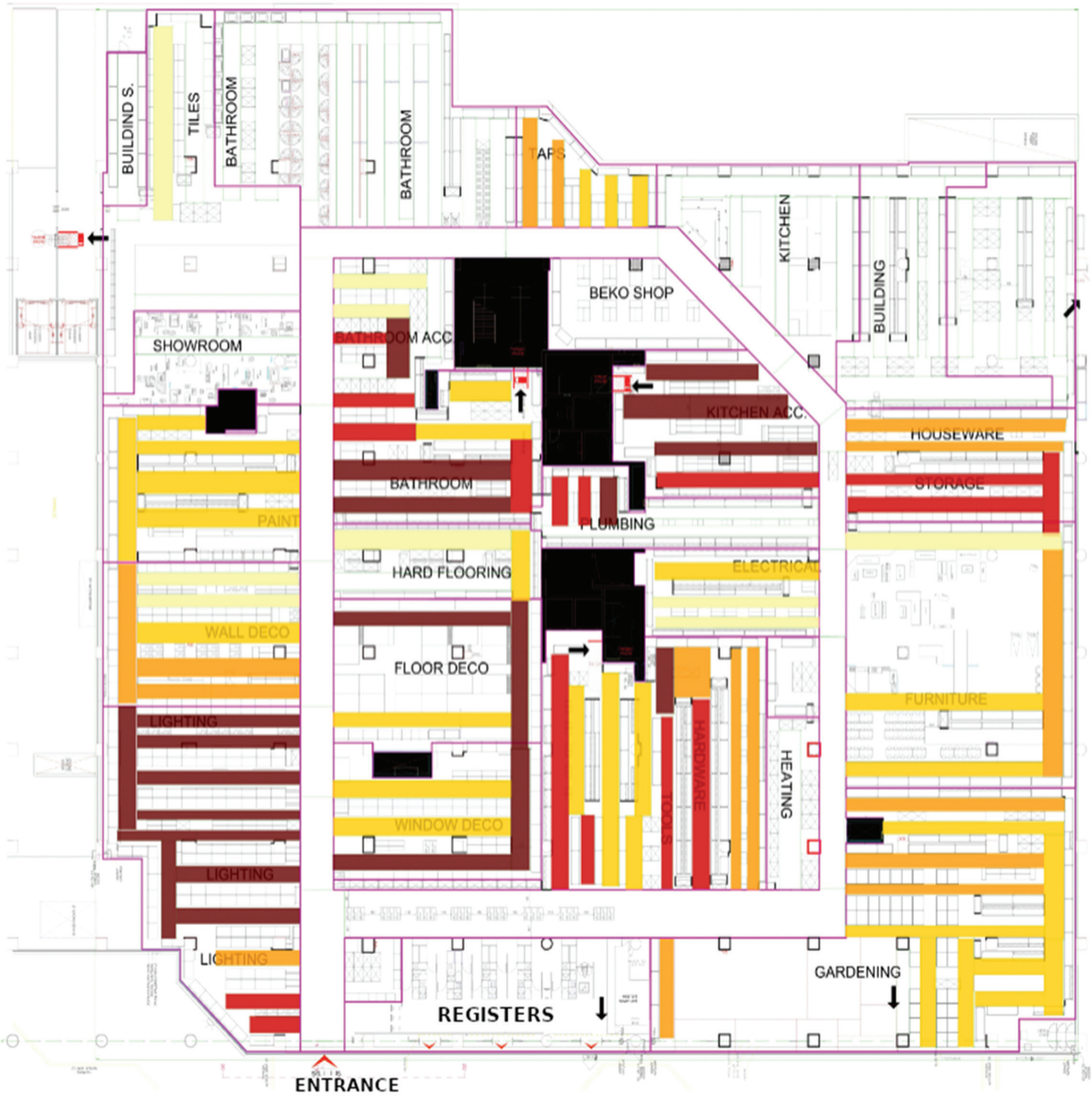


Fig. 7. Heatmap that shows the most visited corridors on 01.06.2024

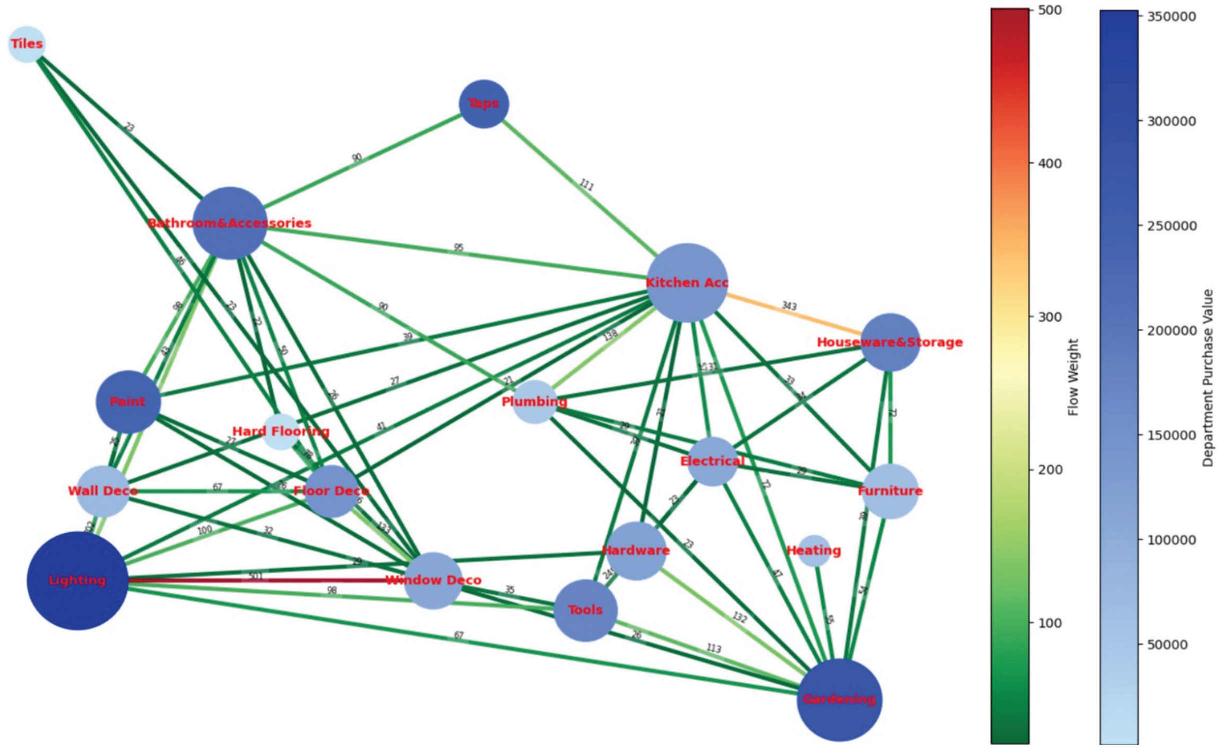


Fig. 8. In the network graph, each node refers to a department and the size of a node corresponds to the total time spent in that department. The darker color refers to a department with a larger total purchase. The edges represent the transitions from one department to another, while the edge color refers to the frequency of that transition

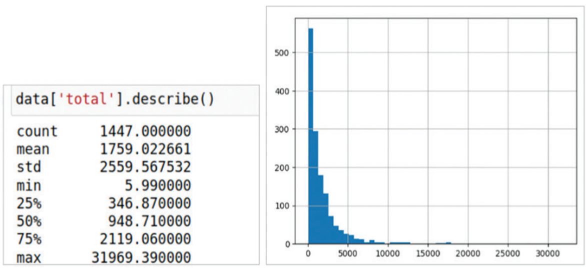


Fig. 9. Total sales data descriptive statistics and the histogram

The dataset was split into 80% for training and 20% for testing, and the classification performance of each method was evaluated in terms of the accuracy under various hyperparameter settings. The results for Model 1 and Model 2 are presented in Table 6 and Table 7, respectively.

Table 6. Model 1 performance comparison of LSTM and XGBoost under different hyper-parameter settings

Model	Configuration	Loss	Accuracy (%)
LSTM1	nl: 1; shl: 128; lr: 0.001; ni: 100; bs: 8	0.79	54.83
LSTM2	nl: 2; shl: 128; lr: 0.001; ni: 100; bs: 8	0.79	54.83
LSTM3	nl: 2; shl: 256; lr: 0.001; ni: 100; bs: 8	0.79	54.83
LSTM4	nl: 2; shl: 512; lr: 0.001; ni: 100; bs: 8	0.79	54.83
XGB1	ne: 2; maxd: 2; lr: 0.1	-	46.21
XGB2	ne: 3; maxd: 2; lr: 0.1	-	46.9
XGB3	ne: 2; maxd: 3; lr: 0.1	-	53.1
XGB4	ne: 3; maxd: 3; lr: 0.1	-	46.21

Table 7. Model 2 performance comparison of LSTM and XGBoost under different hyper-parameter settings

LSTM	Loss	Accuracy (%)	XGBoost	Accuracy (%)
LSTM1	0.79	54.83	XGB1	57.93
LSTM2	0.79	54.83	XGB2	57.24
LSTM3	0.79	54.83	XGB3	56.9
LSTM4	0.8	54.83	XGB4	57.59

The results state that the best LSTM model achieves an accuracy of 54.83% whereas XGB achieves 53.1% for model 1. When it comes to model 2, the results state that XGB achieves a better accuracy than LSTM.

As the initial results were inconclusive, we conducted additional experiments using a modified version of the dataset. Specifically, we removed departments with visit durations of less than 30 seconds in each visit sample. While this adjustment did not reduce the number of samples, it altered the composition of each visit by filtering out short-duration departments. Both algorithms were then applied to each data model.

Additionally, we repeated the experiments after further refining the dataset by excluding departments with visit durations of less than 60 seconds. The results, presented in Table 8, indicate that when using the refined dataset with a 60-second threshold, LSTM models achieved the highest accuracy. This cleaner dataset allowed LSTM models to better capture data characteristics, leading to improved learning. Notably, the loss trajectory of LSTM models on this refined dataset differed significantly from that observed on the original noisy dataset, as illustrated in Fig. 10.

When comparing Model 1 and Model 2 in terms of their impact on algorithm learning, the results suggest that this enhanced data representation did not positively affect classification accuracy. While a richer data representation appeared to improve XGBoost's performance on the noisy dataset, the same effect was not observed when using the cleaned dataset.

Table 8. Accuracy comparison of the algorithms on each of the models using different visit duration thresholds

Model	visit duration<30 removed		visit duration<60 removed	
	Model 1 Acc. (%)	Model 2 Acc. (%)	Model 1 Acc. (%)	Model 2 Acc. (%)
LSTM1	54.83	54.83	58.28*	54.83
LSTM2	54.83	55.86	54.83	54.83
LSTM3	54.83	54.83	56.55	55.17
LSTM4	54.83	54.83	54.83	55.17
XGB1	55.8	56.55	53.79	54.14
XGB2	55.5	56.2	55.17	53.4
XGB3	54.48	55.52	53.45	53.4
XGB4	52.7	54.14	52.76	54.48

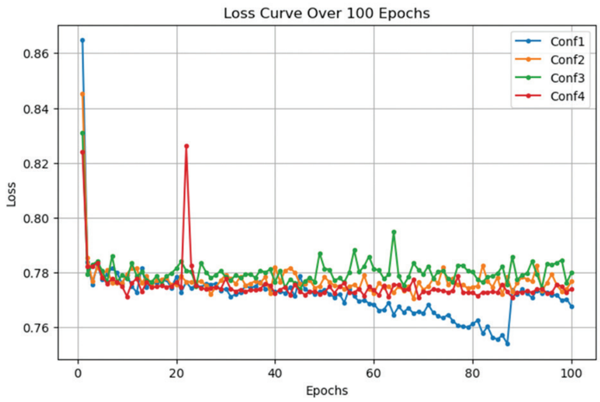


Fig. 10. LSTM loss curve for model 1 when departments with duration <60 secs are ignored

LSTM models appear to have a stronger learning capability than XGBoost for the sequential data used in this study. Throughout the data collection period, a total of 1,447 beacon sessions with valid in-store visit and POS data were gathered. We believe that with a larger dataset, LSTM models could achieve even better performance, further enhancing their ability to capture meaningful sequential patterns.

5. CONCLUSION

In this study, we aimed to observe the in-store behavior of customers at a home improvement retail company. Customer in-store visit data was collected using Bluetooth Low Energy (BLE) beacon devices installed on shelves and shopping carts within the selected store. Cart beacons functioned as receivers, capturing and recording signals from the shelf beacons. Beacons were strategically placed on store shelves to ensure complete coverage, leaving no blind spots. To cover 18 departments spanning approximately 4,800 square meters, 99 beacons were deployed. The duration of stay in each department, the sequence of visits, and the exact visit date and time were recorded in a database. Signals recorded by cart beacons were used to identify the departments where customers spent time, providing valuable insights into customer profiles that can inform marketing decisions.

The correlation analysis revealed a positive relationship between the duration of stay and the purchase amount, particularly in the Floor Deco, Paint, and Taps departments. This information can be leveraged to implement strategic actions, such as expanding these areas or assigning sales assistants to these departments, particularly on specific days of the week. Additionally, we visualized the entire store's data over a given time period using a network diagram. This diagram highlights the departments with the highest sales amounts and longest visit durations while also illustrating customer flow between departments. This visualization provides store managers with a comprehensive overview of in-store behavior, enabling data-driven decisions for optimizing department layouts. Furthermore, insights from customer flow patterns can be leveraged to design targeted marketing campaigns, enhancing the overall shopping experience and sales effectivene

To explore the relationship between in-store behavior and purchase data, we formulated the problem as a multi-class classification task and employed two machine learning models namely, Long Short-Term Memory (LSTM) networks and Extreme Gradient Boosting (XGBoost), for comparative analysis. Experiments were conducted on both the original noisy dataset and a cleaned version, using two distinct data modeling approaches. The first approach utilized only sequential department visit data, while the second model incorporated visit duration information into the sequence representation. The results indicate that both algorithms achieved similar accuracy on the noisy dataset across both data representations, suggesting that adding duration information did not enhance learning. However, when trained on the cleaned dataset, where short-duration department visits were removed, LSTM models demonstrated superior performance, highlighting their ability to better capture meaningful sequential patterns under refined data conditions

These findings underscore the potential of BLE beacon technology for gaining deeper insights into customer behavior, optimizing store layouts, and enhancing personalized marketing strategies. By analyzing in-store movement patterns, retailers can refine product placements, streamline customer journeys, and improve overall shopping experiences. In future work, we plan to collect additional customer visit session data and repeat the experiments with a larger dataset.

6. ACKNOWLEDGEMENTS

This work was supported by the Scientific and Technological Research Council of Türkiye (TÜBİTAK) under grant number 3225040

We sincerely thank Gürsu Gülcü for his invaluable insights and expertise, which played a crucial role in the data analysis process. We also extend our appreciation to İsmail Taha Samed Özkan for his assistance during the data analysis and visualization phase.

7. REFERENCES

[1] F. Zhao, "Analysis of consumer behavior and discussion of personalized marketing strategy in the era of big data", Finance & Economics, Vol. 1, No. 2, 2023.

[2] J. S. Larson, E. T. Bradlow, P. S. Fader, "An exploratory look at supermarket shopping paths", International Journal of Research in Marketing, Vol. 22, No. 4, 2005, pp. 395-414.

[3] E. Haktanır, C. Kahraman, S. C. Onar, B. Öztayşi, S. Cebi, "A state of the art literature review on digital transformation", Intelligent Systems in Digital Transformation: Theory and Applications, Springer, 2022, pp. 3-31.

[4] C. Ke, M. Wu, Y. Chan, K. Lu, "Developing a BLE beacon-based location system using location fingerprint positioning for smart home power management", Energies, Vol. 11, No. 12, 2018, p. 3464.

[5] Q. Liu, X. Yang, L. Deng, "An ibeacon-based location system for smart home control", Sensors, Vol. 18, No. 6, 2018, p. 1897.

[6] P. Spachos, K. N. Plataniotis, "BLE beacons for indoor positioning at an interactive IoT-based smart museum", IEEE Systems Journal, Vol. 14, No. 3, 2020, pp. 3483-3493.

[7] G. Shipkovenski, T. Kalushkov, E. Petkov, V. Angelov, "A beacon-based indoor positioning system for location tracking of patients in a hospital", Proceedings of the 2nd International Congress on Human-Computer Interaction, Optimization and Robotic Applications, Ankara, Turkey, 26-27 June 2020, pp. 1-6.

[8] R. Pangriya, "Beacon technology the future of retail: A review of the literature and swot analysis", A Journal of Management, Vol. 1, No. 11, 2023, pp. 1-11.

[9] D. Jain, M. S. B. Uppal, S. K. Singh, "Exploring Consumers' Readiness And Acceptance Of Beacon Technology", Educational Administration: Theory and Practice, Vol. 30, No. 5, 2024, pp. 3366-3374.

[10] H. Lemsieh, I. Abarar, "Considering the Beacon technology, a Green and eco-friendly revolution in the Moroccan Commercial Malls for a proximity marketing strategy", Proceedings of the E3S Web of Conferences, Istanbul, Turkey, 29-31 October 2024, p. 6.

[11] C. Garcia, S. Inoue, "Relabeling for Indoor Localization Using Stationary Beacons in Nursing Care Facilities", Sensors, Vol. 24, No. 2, 2024, p. 319.

[12] P. Shende, S. Mehendarge, S. Chougule, P. Kulkarni, U. Hatwar, "Innovative ideas to improve shopping mall experience over e-commerce websites using beacon technology and data mining algorithms", Proceedings of the International Conference on Circuit, Power and Computing Technologies, Kollam, India, 20-21 April 2017, pp. 1-5.

[13] L. Zhao, Y. Zuo, K. Yada, "Sequential sequence-to-sequence learning with gated-attention neural networks", Advances in Data Analysis and Classification, Vol. 17, No. 3, 2023, pp. 549-581.

[14] Y. Zuo, A. S. Ali, K. Yada, "Consumer purchasing behavior extraction using statistical learning theory", Procedia Computer Science, Vol. 35, 2014, pp. 1464-1473.

[15] Y. Zuo, K. Yada, E. Kita, "A Bayesian network approach for predicting purchase behavior via direct observation of in-store behavior", Proceedings of the Advanced Methodologies for Bayesian Networks: Second International Workshop, Yokohama, Japan, 16-18 November 2015, pp. 61-75.

[16] L. Li, "Analysis of e-commerce customers' shopping behavior based on data mining and machine learning", Soft Computing, 2023, pp. 1-10.

[17] S. J. Park, C. U. Kang, Y. C. Byun, "Extreme gradient boosting for recommendation system by transforming product classification into regression based on multi-dimensional word2vec", Symmetry, Vol. 13, No. 5, 2021, p. 758.

[18] S. Hochreiter, J. Schmidhuber, "Long short-term memory", Neural Computation, Vol. 9, No. 8, 1997, pp. 1735-1780.

[19] T. Chen, C. Guestrin, "Xgboost: A scalable tree boosting system", Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13-17 August 2016, pp. 785-794.

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