p-ISSN 2521-4209
e-ISSN 2521-4217

جامعة التنمية البشرية
UNIVERSITY OF HUMAN DEVELOPMENT

UHD Journal of Science and Technology

A Scientific periodical issued by University of Human Developement

Vol.9 No.(1) June 2025

2025 2725

www.jst.uhd.edu.iq

Kurdistan Regional Government
University of Human Development

UHD Journal of Science and Technology

A periodic scientific journal issued by University of Human Development

Introduction

UHD Journal of Science and Technology (UHDJST) is a semi-annual journal published by the University of Human Development, Sulaymaniyah, Kurdistan Region, Iraq. UHDJST member of ROAD, e-ISSN: 2521-4217, p-ISSN: 2521-4209 and a member of Crossref, DOI: 10.21928/issn.2521-4217. UHDJST publishes original research in all areas of Science, Engineering, and Technology. UHDJST is a Peer-Reviewed Open Access journal with Creative Commons Attribution Non-Commercial No Derivatives License 4.0 (CC BY-NC-ND 4.0). UHDJST provides immediate, worldwide, barrier-free access to the full text of research articles without requiring a subscription to the journal, and has article processing charge (APC). UHDJST applies the highest standards to everything it does and adopts APA citation/referencing style. UHDJST Section Policy includes three types of publications: Articles, Review Articles, and Letters.

By publishing with us, your research will get the coverage and attention it deserves. Open access and continuous online publication mean your work will be published swiftly, ready to be accessed by anyone, anywhere, at any time. Article Level Metrics allow you to follow the conversations your work has started.

UHDJST publishes works from extensive fields including, but not limited to:

- Pure Science
- Applied Science
- Medicine
- Engineering
- Technology

Scope and Focus

UHD Journal of Science and Technology (UHDJST) publishes original research in all areas of Science and Engineering. UHDJST is a semi-annual journal published by the University of Human Development, Sulaymaniyah, Kurdistan Region, Iraq. We believe that if your research is scientifically valid and technically sound then it deserves to be published and made accessible to the research community. UHDJST aims to provide a service to the international scientific community enhancing swap space to share, promote and disseminate the academic scientific production from research applied to Science, Engineering, and Technology.

SEARCHING FOR PLAGIARISM

Plagiarism Policy: The UHD Journal of Science and Technology is committed to upholding the highest standards of academic integrity and originality. As such, we do not tolerate plagiarism in any form. Authors submitting manuscripts to the journal must ensure that their work is entirely original and properly cited. Plagiarism, including self-plagiarism and the use of verbatim text from other sources without appropriate attribution, will result in immediate rejection of the submission. To maintain the integrity of scholarly research, all submissions undergo rigorous plagiarism checks using industry-standard software. Authors found to have engaged in plagiarism, at any stage before publication of the manuscript - before or after acceptance, during editing or at page proof

stage, will be barred from submitting to the journal in the future. By submitting their work to the UHD Journal of Science and Technology, authors affirm that their manuscript is their own original work and that any borrowed content is properly cited.

Section Policies

No.	Title	Peer Reviewed	Indexed	Open Submission
1	Articles: This is the main type of publication that UHDJST will produce	☑	☑	☑
2	Review Articles: Critical, constructive analysis of the literature in a specific field through summary, classification, analysis, comparison.	☑	☑	☑
3	Letters: Short reports of original research focused on an outstanding finding whose importance means that it will be of interest to scientists in other fields.	☑	☑	☑

PEER REVIEW POLICIES

At UHDJST we are committed to prompt quality scientific work with local and global impacts. To maintain a high-quality publication, all submissions undergo a rigorous review process. Characteristics of the peer review process are as follows:

- The journal peer review process is a "double-blind peer review".
- Simultaneous submissions of the same manuscript to different journals will not be tolerated.
- Manuscripts with contents outside the scope will not be considered for review.
- Papers will be refereed by at least 2 experts as suggested by the editorial board.
- In addition, Editors will have the option of seeking additional reviews when needed. Authors will be informed when Editors decide further review is required.
- All publication decisions are made by the journal's Editors-in-Chief on the basis of the referees' reports. Authors of papers that are not accepted are notified promptly.
- All submitted manuscripts are treated as confidential documents. We expect our Board of Reviewing Editors, Associate Editors and reviewers to treat manuscripts as confidential material as well.
- Editors, Associate Editors, and reviewers involved in the review process should disclose conflicts of interest resulting from direct competitive, collaborative, or other relationships with any of the authors, and remove oneself from cases in which such conflicts preclude an objective evaluation. Privileged information or ideas that are obtained through peer review must not be used for competitive gain.
- Our peer review process is confidential and the identities of reviewers cannot be revealed.

Note: UHDJST is a member of CrossRef and CrossRef services, e.g., CrossCheck. All manuscripts submitted will be checked for plagiarism (copying text or results from other sources) and self-plagiarism (duplicating substantial parts of authors' own published work without giving the appropriate references) using the CrossCheck database. Plagiarism is not tolerated.

For more information about CrossCheck/iThenticate, please visit http://www.crossref.org/crosscheck.html.

OPEN ACCESS POLICY

This journal provides immediate open access to its content on the principle that making research freely available to the public supports a greater global exchange of knowledge. Open Access (OA) stands for unrestricted access and unrestricted reuse which means making research publications freely available online. It access ensures that your work reaches the widest possible audience and that your fellow researchers can use and share it easily. The mission of the UHDJST is to improve the culture of scientific publications by supporting bright minds in science and public engagement.

UHDJST's open access articles are published under a Creative Commons Attribution CC-BY-NC-ND 4.0 license. This license lets you retain copyright and others may not use the material for commercial purposes. Commercial use is one primarily intended for commercial advantage or monetary compensation. If others remix, transform or build upon the material, they may not distribute the modified material. The main output of research, in general, is new ideas and knowledge, which the UHDJST peer-review policy allows publishing as high-quality, peer-reviewed research articles. The UHDJST believes that maximizing the distribution of these publications - by providing free, online access - is the most effective way of ensuring that the research we fund can be accessed, read and built upon. In turn, this will foster a richer research culture and cultivate good research ethics as well. The UHDJST, therefore, supports unrestricted access to the published materials on its main website as a fundamental part of its mission and a global academic community benefit to be encouraged wherever possible.

Specifically:

- The University of Human Development supports the principles and objectives of Open Access and Open Science
- UHDJST expects authors of research papers, and manuscripts to maximize the opportunities to make their results available for free access on its final peer-reviewed paper
- All manuscript will be made open access online soon after final stage peer-review finalized.
- This policy will be effective from 17th May 2017 and will be reviewed during the first year of operation.
- Open Access route is available at http://journals.uhd.edu.iq/index.php/uhdjst for publishing and archiving all accepted papers,
- Specific details of how authors of research articles are required to comply with this policy can be found in the Guide to Authors.

ARCHIVING

This journal utilizes the LOCKSS and CLOCKSS systems to create a distributed archiving system among participating libraries and permits those libraries to create permanent archives of the journal for purposes of preservation and restoration.

LOCKSS: Open Journal Systems supports the LOCKSS (Lots of Copies Keep Stuff Safe) system to ensure a secure and permanent archive for the journal. LOCKSS is open source software developed at Stanford University Library that enables libraries to preserve selected web journals by regularly polling registered journal websites for newly published content and archiving it. Each archive is continually validated against other library caches, and if the content is found to be corrupted or lost, the other caches or the journal is used to restore it.

CLOCKSS: Open Journal Systems also supports the CLOCKSS (Controlled Lots of Copies Keep Stuff Safe) system to ensure a secure and permanent archive for the journal. CLOCKSS is based upon the open-source LOCKSS software developed at Stanford University Library that enables libraries to preserve selected web journals by regularly polling registered journal websites for newly published content and archiving it. Each archive is continually validated against other library caches, and if the content is found to be corrupted or lost, the other caches or the journal is used to restore it.

PUBLICATION ETHICS

Publication Ethics and Publication Malpractice Statement

The publication of an article in the peer-reviewed journal UHDJST is to support the standard and respected knowledge transfer network. Our publication ethics and publication malpractice statement is mainly based on the Code of Conduct and Best-Practice Guidelines for Journal Editors (Committee on Publication Ethics, 2011) that includes;

- General duties and responsibilities of editors.
- Relations with readers.
- Relations with the authors.
- Relations with editors.
- Relations with editorial board members.
- Relations with journal owners and publishers.
- Editorial and peer review processes.
- Protecting individual data.
- Encouraging ethical research (e.g. research involving humans or animals).
- Dealing with possible misconduct.
- Ensuring the integrity of the academic record.
- Intellectual property.
- Encouraging debate.
- Complaints.
- Conflicts of interest.

ANIMAL RESEARCHES

- For research conducted on regulated animals (which includes all live vertebrates and/or higher invertebrates), appropriate approval must have been obtained according to either international or local laws and regulations. Before conducting the research, approval must have been obtained from the relevant body (in most cases an Institutional Review Board, or Ethics Committee). The authors must provide an ethics statement as part of their Methods section detailing full information as to their approval (including the name of the granting organization, and the approval reference numbers). If an approval reference number is not provided, written approval must be provided as a confidential supplemental information file. Research on non-human primates is subject to specific guidelines from the Weather all (2006) report (The Use of Non-Human Primates in Research).
- For research conducted on non-regulated animals, a statement should be made as to why ethical approval was not required.
- Experimental animals should have been handled according to the highest standards dictated by the author's institution.
- We strongly encourage all authors to comply with the '*Animal Research: Reporting In Vivo Experiments*' (ARRIVE) guidelines, developed by NC3Rs.
- Articles should be specific in descriptions of the organism(s) used in the study. The description should indicate strain names when known.

ARTICLE PROCESSING CHARGES

UHDJST is an Open Access Journal (OAJ) and has article processing charges (APCs). The published articles can be downloaded freely without a barrier of admission.

Address

University of Human Development, Sulaymaniyah-Kurdistan Region/Iraq
PO Box: Sulaymaniyah 6/0778

Contact

Principal Contact

Dr. Aso Darwesh

Editor-in-Chief

University of Human Development – Sulaymaniyah, Iraq

Email: jst@uhd.edu.iq

Support Contact

UHD Technical Support

Phone: +964 773 393 5959

Email: jst@uhd.edu.iq

Contents

ORIGINAL RESEARCH ARTICLE UHD JOURNAL OF SCIENCE AND TECHNOLOGY

An Image Analysis for Designing an Optimal Stirrer in Metal Matrix Composites Manufacturing

Farooq Muhammad, Prof. Dr. Muzhir Shaban Al-Ani, Hamsa D. Majeed

Department of Information Technology, College of Science and Technology, University of Human Development, Kurdistan Region, Iraq

ABSTRACT

The global market for aluminum-based composites, widely used in manufacturing and construction, is expected to grow significantly. However, enhancing the cost-to-performance ratio is essential to improving their commercial viability. Efficient mixing plays a critical role in many industrial and chemical applications. Stir casting is the leading method for producing aluminum alloy matrix composites, but achieving a uniform particle distribution remains a significant challenge. In this study, the optimal stirrer design was identified using image processing techniques to analyze the distribution of ceramic grains. The stirrer that achieved the most uniform grain distribution was selected, eliminating the need for destructive testing. The mechanical properties of the final products validated the accuracy of the image analysis results.

Index Terms: Aluminum Alloy Matrix Composites, Cost-to-performance ratio, Mixing Performance, Stir Casting, Ceramic Reinforcement Particles, Stirrer Design, Image Processing, Mechanical Properties

1. INTRODUCTION

As energy shortages and environmental pollution risks increase, the demand for stronger, lighter, and more environmentally friendly materials continues to grow. Aluminum alloy ceramic matrix composites are extensively used across various industries and are expected to expand further as efforts to protect the environment and reduce fossil fuel usage intensify. Stir casting, also known as the vortex process, is widely recognized as an economical method for producing metal matrix composites. However, achieving mixing homogeneity remains a significant challenge, heavily influenced by stirrer design. In this study, the authors isolated the effect of stirrer design by keeping other key variables – stirring time, stirring speed, and stirrer position – constant, thereby eliminating potential interference or interaction effects from these factors. These parameters were fixed as follows: Stirring time at 60 s, stirring speed at 240 rpm, and stirrer position at two-thirds of the crucible height. This approach allowed the exclusive evaluation of the stirrer design's function. Fig. 1 provides a flowchart summarizing the research methodology used to achieve the study's objectives.

Using image processing to assess the uniform distribution of solid particles in liquids at elevated temperatures, before solidification. Accurately measuring particle distribution in high-temperature fluids presents considerable difficulties due to several key challenges. These include the requirement for sensors that can withstand high temperatures, real-time, non-destructive measurement techniques capable of providing sufficient penetration depth, and the ability to monitor dynamic changes in particle distribution amidst stirring and thermal fluctuations. Traditional methods such as microscopy are inadequate due to their destructive nature.

Access this article online

DOI: 10.21928/uhdjst.v9n1y2025.pp1-9

E-ISSN: 2521-4217
P-ISSN: 2521-4209

Corresponding author's e-mail: Farooq Muhammad, College of Science and Technology, University of Human Development, Kurdistan Region, Iraq. E-mail: farooq.muhammad@uhd.edu.iq

Received: 19-10-2024 Accepted: 28-12-2024 Published: 11-01-2025

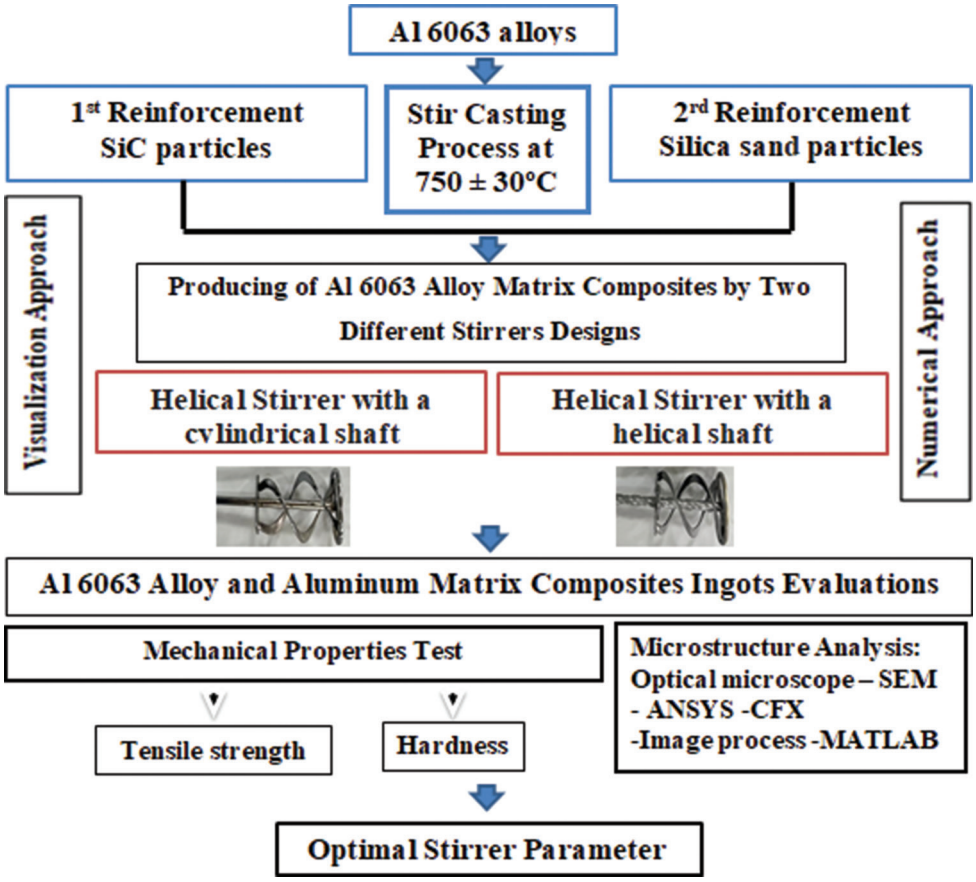


Fig. 1. Diagram of the research work.

Potential alternatives, including high-temperature ultrasound, electromagnetic methods, and specialized optical techniques, each possess their limitations and drawbacks. As such, image processing emerges as a promising alternative deserving further exploration and evaluation.

2. RELATED WORK

Digital image processing (DIP) technology has significant potential in particle mixture analysis. Nowadays, surface analysis has many applications across various fields, and one important area is material surface analysis [1]. Proposed a model where a color histogram was implemented to analyze particle mixtures in fluidized beds. The proposed method serves as a proof of concept, demonstrating that mixing particles changes the color distribution of the image and can be used to evaluate the quality of the mixing process. The results show that particle mixing alters the histograms of images taken at different times, providing an interesting avenue for exploring automatic particle mixing techniques based on image processing. On the other hand [2], [3], proposed an approach in which texture-based image processing plays a more prominent role than standard image enhancement. This method facilitates the detection of small obstacles during surface preparation by making individual features and surface details more visible. Various algorithms and mathematical models were applied to compare different image surfaces. In addition, statistical measures were used to evaluate the performance of this approach, further validating its effectiveness in surface analysis [4] studied the effects of adding micro, nano, and hybrid particles in a 1:1 ratio of Al2O3 and TiO2 to epoxy resin on thermal conductivity. The study analyzed samples before and after immersion in 0.3 N HCl acid for 14 days, considering weight fractions of 0.02, 0.04, 0.06, and 0.08, with a thickness of 6 mm. Image analysis techniques were employed to evaluate the homogeneity of the mixtures, highlighting the practical application of DIP in material science.

An analysis and comparison of the image analysis approach with other existing methods are necessary to evaluate the homogeneity of mixing in a moving casting. The selection

of an appropriate method depends on several factors. These include particle properties such as size, shape, density, and concentration; fluid properties such as viscosity, density, and surface tension; type of mixing equipment, including mixer type, design, and operating conditions; required level of accuracy and sensitivity; and available resources and expertise. This research focuses on finding non-destructive tests. Even if we assume that destructive tests are acceptable, microscopic analysis, as one of the destructive tests, has limitations. It is time-consuming and provides data only from limited sample areas. It is time-consuming and provides data only from limited sample regions [5].

3. METAL MATRIX COMPOSITES

In 2022, the market for metal matrix composites (MMCs) was valued at USD 211.3 billion. It is projected to grow significantly, reaching USD 369.3 billion by 2032, with an annual growth rate of 6.40% during the forecast period from 2023 to 2032. The increasing demand for lightweight materials in the automotive and aerospace sectors is expected to be a key driver of this market growth. Fig. 2 illustrates the projected trend for MMCs in the coming years.

Aluminum matrix composites (AMCs) use pure aluminum or an alloy as the matrix and are becoming more popular in industrial applications due to their outstanding mechanical and tribological properties. The properties of Al alloys might be greatly tailored by adding ceramic reinforcing particles by stir casting. AA 6063 ingot was chosen as a matrix alloy which was analyzed carefully by using an X-MET8000 Handheld XRF analyzer and the compositions of the used material are shown in Tables 1-4.

In previous work [11], the visualization procedure was performed using two stirrer designs: A helical stirrer with a cylindrical shaft (D4) and a helical stirrer with a helical shaft (D5), as shown in Fig. 3. A glass jar, similar in dimensions to a graphite crucible, was filled with fluids of varying viscosities – water (low viscosity) and engine oil (high viscosity) – to simulate the stirring process of molten aluminum and SiC or silica particles. The jar had a height of 150 mm and a diameter of 100 mm. A 1% volume fraction of real micro-sized SiC and silica particles was added to the fluid, and stirring was conducted using an SEI-WA MG-915 radial drill connected to different stirrer designs. The stirring speed was set at 240 rpm, with the stirrer positioned at 30% of the crucible's height from the base. Images were captured immediately after removing the stirrer following a constant holding time of 60 s of stirring. The images, taken from the open top of the jar, minimized glare issues caused by the glass. This visualization experiment was repeated under consistent conditions for all five stirrer designs: Single-blade, double-blade, multi-stage stirrer, helical stirrer, and helical stirrer with a helical shaft, as shown in Figs. 4 and 5. Both stirrers (D4 and D5) were used to mix a 0.05% volume fraction of SiC particles and silica sand particles (50–80 microns in size) under two conditions: Once with one liter of water and then with one liter of engine oil. The mixing was conducted for 60 s at a stirring speed of 250 rpm, with the stirrer operating within the top two-thirds of the glass jar's height. Images were captured from the open top portion, as shown in Fig. 6, to ensure clarity

Existing Methods	Description	Limitations	References
Visual Inspection	Observing the distribution of particles with the naked eye or using optical devices.	Subjective and lacks quantitative accuracy.	[6]
Conductivity Measurements	Measuring the electrical conductivity of the mixture, which can change with particle concentration.	Limited Sensitivity to Certain Components: They are primarily effective for detecting ionic species and may not provide accurate results for non-ionic or poorly conductive components in the mixture.	[7]
Computational Fluid Dynamics (CFD)	Using CFD models to predict the flow behavior and particle distribution within the mixing vessel.	High costs in chaotic flows, impairing accurate evaluation of mixing and homogeneity.	[8]
Optical Techniques, Laser Diffraction	Measuring the particle size distribution and identifying any significant variations within the mixture.	Laser diffraction assumes spherical particles, leading to inaccurate size measurements and hindering homogeneity assessment in mixtures with irregularly shaped particles.	[9]
Ultrasonic Testing	Using ultrasound waves to measure the acoustic properties of the mixture, which can be affected by particle distribution.	Anisotropic materials have varying ultrasonic wave velocities due to structural variations, complicating accurate characterization.	[10]

TABLE 1: Composition of the AA 6063 ingot

Element	Al	Si	Fe	Zr	Mn	Mg	Zn	Ni	Pb	Sn	Ti
Average content	98.76	0.19	0.16	0.02	0.02	0.74	<0.00	<0.00	0.04	0.07	<0.01

TABLE 2: Chemical Composition of Silicon Carbide (SiC)

Component	SiC	C	Si	SiO_2	Fe_2O_3
Wt.%	97.2	0.28	0.23	1.73	0.56

TABLE 3: Properties of Silicon Carbide (SiC)

Properties	Measure
Melting point	2200 to 2700
Density	3.1(g/cm^3)
Coefficient of Thermal Expansion	4.1(μm/m/°C)
Fracture toughness	4.6(MPa-m1/2)
Poisson's ratio	0.14
Color	Black

TABLE 4: Properties of (Sio_2)*

Constituent	Percentage
Sio_2	89
Al_2O_3	3.75
Fe_2O_3	1.94
MgO	1.02
Cao	0.99
Loss on ignition	2.63

*Silica sand samples from the Mass company (Iraq) (50–80-micron size) were used. The chemical composition is given in Table 4.

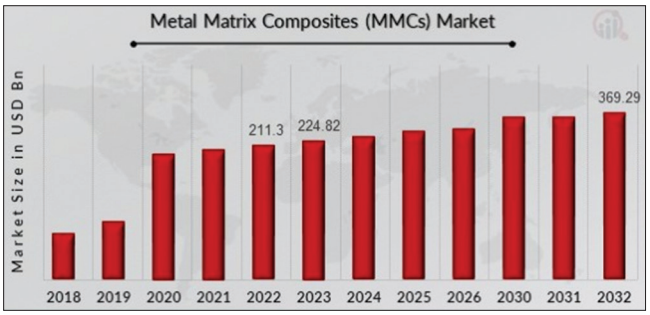


Fig. 2. MMC trend.

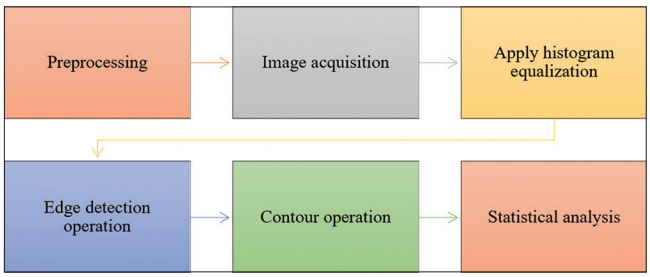


Fig. 3. Implemented approach.

and avoid glare issues. These images were then analyzed using MATLAB to evaluate the mixing process.

5. IMAGE PROCESSING ANALYSIS

Traditional image processing has been widely utilized to enhance the aluminum alloy casting process by automating quality control and optimizing casting methods. Various image processing techniques have been employed for aluminum alloy analysis, including X-ray CT for internal defect detection, surface inspection for identifying surface flaws, real-time monitoring for process optimization, microstructure analysis for understanding material properties, dimensional measurement for precision, and robotic vision for automated handling [12], [13]. Mixing is a critical process that directly influences the quality of final products in several industries. Ensuring homogeneous mixtures requires a thorough examination of the mixing process [14]. This study highlights an alternative, non-invasive method for analyzing particle mixing and segregation using image processing, which has shown great potential. A proposal to utilize low-cost cameras combined with image processing has been introduced as an effective, affordable solution [15]. High-tech methods for analyzing mixing processes are often costly and demand strict safety protocols, making image processing a promising alternative.

In recent years, image analysis has gained significance due to the increasing volume of images transmitted through various media in daily life. Image processing involves applying specific operations to enhance images or extract meaningful information from them. An image is mathematically represented as a two-dimensional function $(x,)$, where x and y are spatial coordinates, and the amplitude at any coordinate pair corresponds to the image intensity. When, the amplitude values of f are finite and discrete, the image is termed a digital image [15]. Digital images are composed of pixels, each with a specific location and value. This paper uses statistical methods to identify optimal mixtures by analyzing images captured under varying incident light conditions and with different stirrer designs. Image content analysis is an essential step for producing accurate and unbiased evaluations.

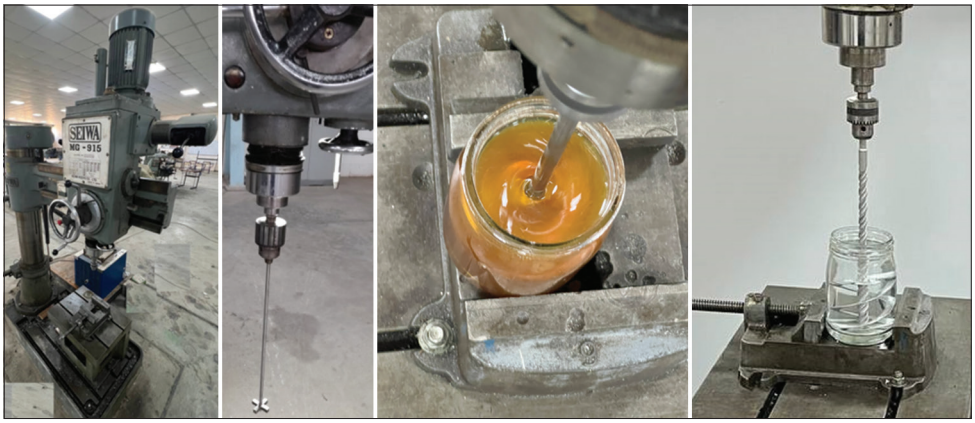


Fig. 4. SEIWA MG-915 Radial Drill setting.

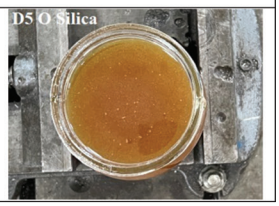
Stirrer types		After 60 seconds stirring at (240 rpm)	
		Water + SiC particles	Oil + Silica particles
Helical stirrer with a cylindrical shaft (D4)			
Helical stirrer with a helical shaft (D5)			

Fig. 5. Images were captured from the exposed upper section.

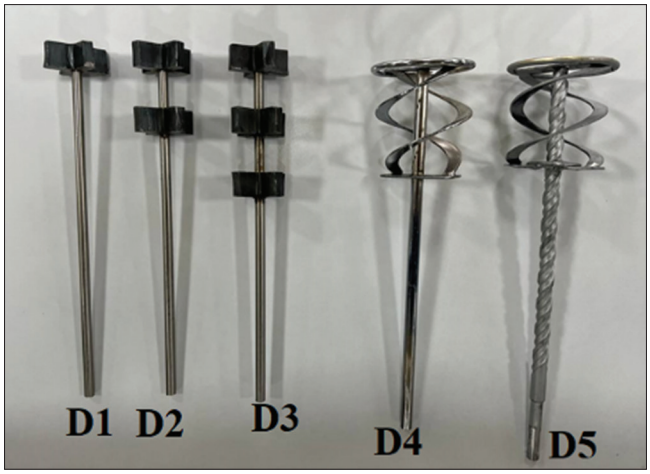


Fig. 6. Five different designs of stirrers, a four blades single stirrer (D1), double stages stirrer (D2), a multi-stages stirrer (D3), a helical stirrer (D4), and a helical stirrer with helical shaft (D5)

Fig. 7. Helical Stirrer helical stirrer with a cylindrical shaft (D4) and a helical stirrer with a helical shaft (D5).

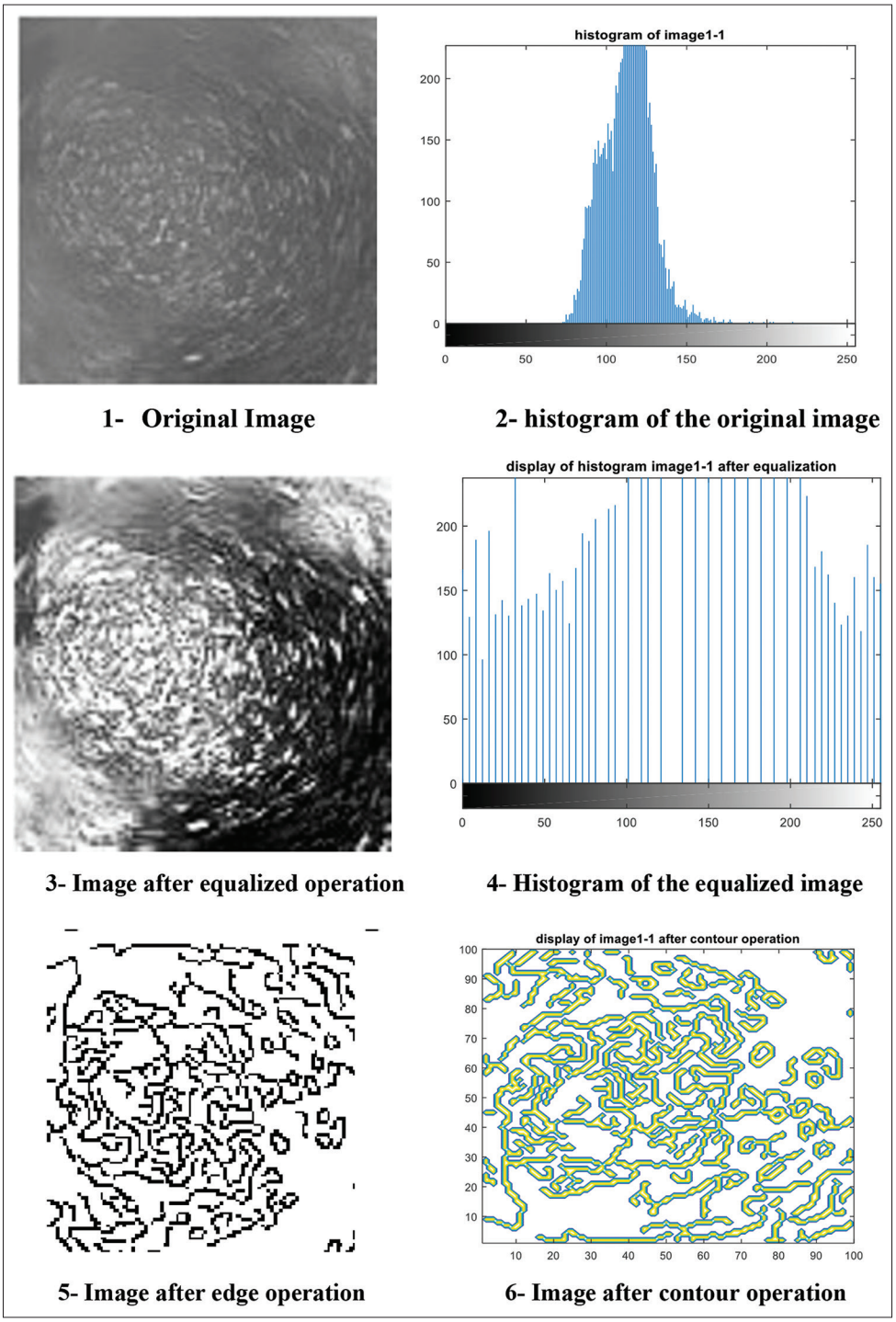


Fig. 8. Image process steps from original to after-edge operations.

Algorithms and techniques for processing color information in three-dimensional scenes and converting them into two-dimensional images in trichromatic color models and digital color spaces demonstrate similarities between color and grayscale image processing [16], [17]. Image processing is integral to modern manufacturing, enabling the manipulation and analysis of digital images to enhance product quality, extract critical information, and optimize processes. Techniques such as object recognition, segmentation, and optimization improve efficiency, automate quality control,

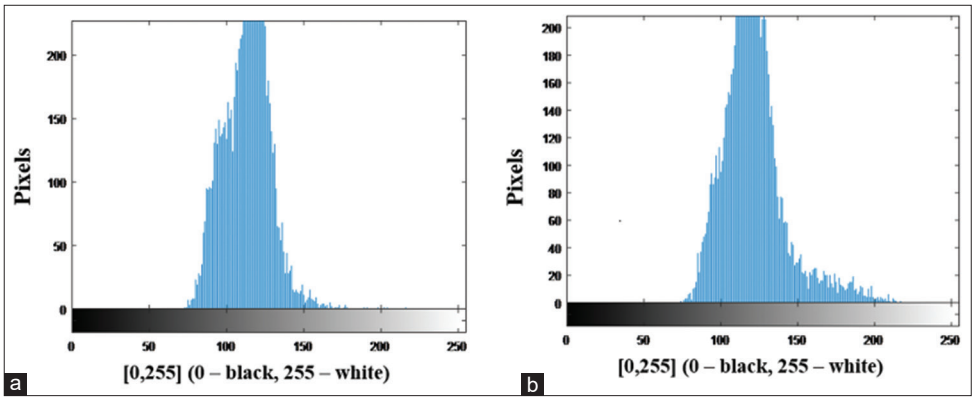


Fig. 9. (a and b) D4 stirrer mixing performance.

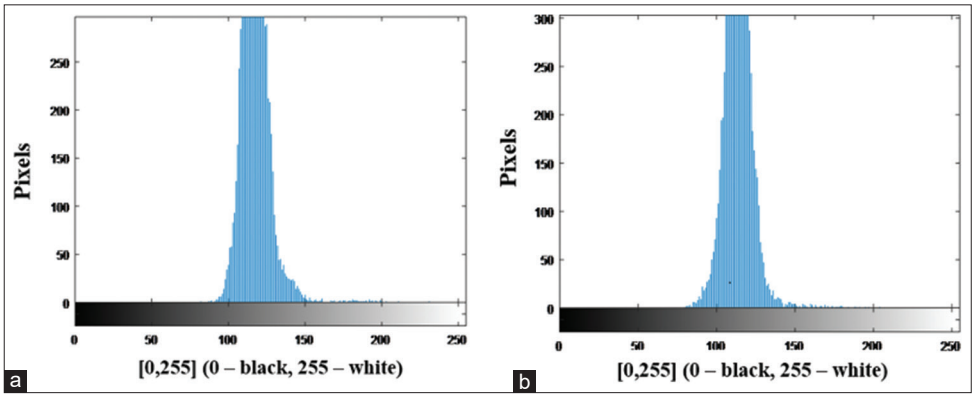


Fig. 10. (a and b) D5 stirrer mixing performance.

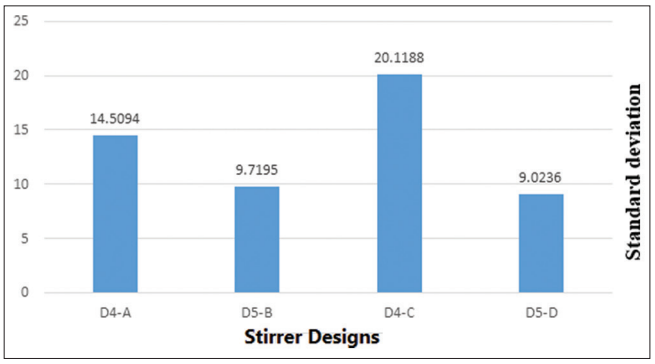


Fig. 11. The standard deviation for D4 and D5 stirrers.

detect defects, and streamline diagnostics, contributing to reliable and cost-effective production systems [18].

Image analysis involves extracting valuable information from images, typically using two main approaches: Spatial domain and frequency domain methods [19].

The statistical approach belongs to the spatial domain category, assessing various functions based on system objectives. Statistical measures such as mean and standard deviation play a significant role in image analysis by providing insights into the properties of incident light and mixing quality [20]. In this study, after completing the sample mixing process, images of the mixture were captured and analyzed using digital image processing techniques. The process involves several steps: Pre-processing, image acquisition, histogram equalization, edge detection, and contour detection. These steps are illustrated in Fig. 10.

Image processing provides an affordable alternative. It is a type of signal processing that generates digital images as its output or a set of properties or parameters connected to the image [15]. This is accomplished by the analysis of the pictures obtained during the mixing process by the camera. As a result of the mixing procedure, it has been discovered that the colors in the image do change. Image analysis was divided into two sections:

5.1. Data Required

In the practical part of this study, four cases were developed as follows:

Case 1 (image1_1): Stirring Silicon carbide microparticles in one liter of water using a helical stirrer with a cylindrical shaft for 60 s at 240 rpm under laboratory conditions.

Case 2 (image1_2): Stirring Silicon carbide microparticles in one liter of water using a helical stirrer with a helical shaft for 60 s at 240 rpm under laboratory conditions.

Case 3 (image2_1): Stirring Silicon carbide microparticles in one liter of engine oil (high viscosity) using a helical stirrer with a cylindrical shaft for 60 s at 240 rpm under laboratory conditions.

Case 4 (image2_2): Stirring Silicon carbide microparticles in one liter of engine oil (high viscosity) using a helical stirrer with a helical shaft for 60 s at 240 rpm under laboratory conditions.

5.2. Implemented Approach

After completing the sample mixing, an image of the mixture is captured and analyzed based on the principles of digital image processing. The processing stage involves several steps, which are illustrated in Fig. 7:

1. Pre-processing: The mixture is prepared, and an image is captured. A portion of the image, sized 100 × 100 pixels, is selected for further processing.
2. Image Acquisition: The captured image is formatted and prepared for the subsequent processing stages.
3. Histogram Equalization: Histogram equalization is applied to enhance image clarity and improve visibility for the next steps.
4. Edge Detection: This step involves identifying the edges within the image using the canny operator, which provides superior edge detection compared to other methods.
5. Contour Detection: The contours of shapes in the image are identified to outline the structures present.
6. Statistical Analysis: Various statistical measures are applied to analyze the image properties, including:
    a. Standard Deviation: Measures the dispersion of data points and their deviation from the arithmetic mean, indicating the spread of pixel values.
    b. Arithmetic Mean: Represents a measure of central tendency, summarizing the data set with a single value that denotes its average distribution.
    c. Entropy: Quantifies the randomness or texture of the input image, providing insights into its complexity and information content.

After processing images as shown its the steps in Fig. 8, the performances of the stirrers were calculated according to their pixel distribution.

The findings from Figs. 9 and 10 suggest that D5 consistently outperforms D4 across various conditions. This includes low and high viscosity (water or engine oil) and low and high ceramic density (Silica sand or SiC). The mixing performance of D4 is characterized by a pixel distribution ranging from 75 to 165 grayscale values, while D5 shows a more refined pixel distribution between 90 and 150 grayscale values. This difference indicates that D5 achieves a higher level of homogeneity in the mixing process, as it maintains a more consistent pixel range, suggesting better uniformity in the particle distribution.

Fig. 11 illustrates the distribution of standard deviation values for the processed images. Standard deviation is a measure of the variation or spread within a set of values. A low standard deviation indicates that the pixel intensities are close to the average, while a high standard deviation suggests a greater deviation from the average. In all conditions, the standard deviation values for D4 are higher than those for D5. This suggests that D4 exhibits greater variation in pixel intensities, indicating a less homogeneous mixing process compared to D5, which shows more consistent pixel values and a higher degree of homogeneity.

6. CONCLUSION

1. This study investigates the use of image processing analysis techniques to optimize stirrer design for the production of aluminum alloy composites. The study involved testing five different stirrer designs and two types of ceramic particle reinforcement using the stir-casting method.
2. The image processing steps included pre-processing, image acquisition, histogram equalization, edge detection, contour operation, and statistical analysis.
3. Image analysis techniques played a crucial role in determining the most effective stirrer design by analyzing the homogeneity of the mixture.
4. The accuracy of the image analysis findings was validated by evaluating the mechanical properties of the produced composites.
5. The analysis of pixel distribution revealed that D5 consistently outperformed D4 under various conditions.
6. Standard deviation values for D4 were higher than those for D5, indicating greater variation in pixel intensity and

suggesting less homogeneous mixing in D4.

7. The helical stirrer with a helical shaft (D5) demonstrated superior particle mixing and improved homogeneity compared to the cylindrical shaft stirrer (D4).

7. ACKNOWLEDGMENT

Sincere appreciation is extended to all individuals who provided support to this paper. Their invaluable guidance, assistance, and encouragement are deeply acknowledged and greatly appreciated.

8. DECLARATION OF CONFLICTING INTERESTS

The authors stated that there are no potential conflicts of interest related to the research, writing, or publication of this article.

9. FUNDING

The authors did not receive financial support for any aspect of the research, authorship, or publication of this article.

10. DATA AVAILABILITY

No new data were generated or analyzed in support of this research.

REFERENCES

[1] S. A. Mohd Zuki, N. Abdul Rahman and I. Mohd Yassin. "Particle mixing analysis using digital image processing technique". *Journal of Applied Sciences*, vol. 14, no. 13, pp. 1391-1396, 2014.

[2] A. A. Z. Hameed, F. Hammad Anter and M. Shaban Al-Ani. "Study of effect acidic solution (HCl) and (EP/Al 2 O 3 & EP/TiO_2) hybrid on thermal conductivity of epoxy resin". *Iraqi Journal of Physics*, vol. 15, pp. 92-99, 2017.

[3] A. A. Z. Hameed, F. Hammad Anter, and M. Shaban Al-Ani. "Composite material surface analysis based image texture analysis". *International Journal of Computer Applications*, vol. 129, pp. 22-26, 2015.

[4] A. A. Z. Hameed, M. Shaban Al-Ani and F. Hammad Anter. "The effects of nano alumina: A comparative study of material surface texture analysis". *International Journal of Business*, vol. 1, no. 4, pp. 26-33, 2015.

[5] D. Bellato, I. P. Marzano and P. Simonini. "Microstructural analyses of a stabilized sand by a deep-mixing method". *Journal of Geotechnical and Geoenvironmental Engineering*, vol. 146, p. 04020032, 2020.

[6] P. Suetens and A. Oosterlinck. Critical review of visual inspection. *Digital Image Processing, Society of Photo-Optical Instrumentation Engineers*, Vol. 528, pp. 240-252, 1985.

[7] F. Opekar, K. Štulík and M. Fišarová. "Some possibilities and limitations of contactless impedimetric determinations of organic liquids in aqueous solutions. The interference from ionic compounds". *Electroanalysis*, vol. 22, no. 20, pp. 2353-2358, 2010.

[8] K. Li, C. Savari, and M. Barigou. "Predicting complex multicomponent particle-liquid flow in a mechanically agitated vessel via machine learning". *Physics of Fluids*, vol. 35, no. 5, p. 053301 2023.

[9] S. Andrews, D. Nover and S. G. Schladow. "Using laser diffraction data to obtain accurate particle size distributions: The role of particle composition". *Limnology and Oceanography: Methods*, vol. 8, no. 10, pp. 507-526, 2010.

[10] Y. Pamungkas. "Ultrasonic non-destructive materials characterization". In: *Non-Destructive Materials Characterization and Evaluation. Springer Series in Materials Science*. Ch. 2. Springer, Berlin, Heidelberg, 2023.

[11] F. Muhammad and S. Jalal. "Optimization of stirrer parameters by Taguchi method for a better ceramic particle stirring performance in the production of aluminum alloy matrix composite". *Cogent Engineering*, vol. 10, no. 1, p. 2154005, 2023.

[12] T. F. Stocker, J. Mehringer, H. Frechen, F. Sukowski, F. Schäfer and D. Freier. "Reduction of rejects by combining data from the casting process and automatic X-ray inspection". *E-Journal of Nondestructive Testing*, vol. 29, p. 6, 2024.

[13] N. Muralidhar. "Comparative study of image processing algorithms to detect defects in cast components". *DBS Business Review*, vol. 5, pp. 59-73, 2023.

[14] N. A. Rahman, S. Akma and M. Zuki. "A Review of Image Processing Technique in Particle Mixing". In: *2012 IEEE 8th International Colloquium on Signal Processing and its Applications*. IEEE, Malacca, Malaysia, pp. 466-469, 2012.

[15] R. S. Dwivedi. "*Remote Sensing of Soils*". Springer, Berlin, Germany, 2017.

[16] M. Ravindranath and C. S. Sastry. "Compressed Sensing for Reconstruction of Reflectance Spectra from Tristimulus Values". In: *Proceedings - 2nd Vaagdevi International Conference on Information Technology for Real World Problems, VCON 2010*. IEEE, Warangal, India, pp. 79-82, 2010.

[17] M. S. Al-Ani and K. M. Ali Alheeti. "Precision statistical analysis of images based on brightness distribution". *Advances in Science, Technology and Engineering Systems*, vol. 2, no. 4, pp. 99-104, 2017.

[18] S. Boopathi, B. K. Pandey, and D. Pandey. "Advances in artificial intelligence for image processing: Techniques, applications, and optimization". In: *Handbook of Research on Thrust Technologies' Effect on Image Processing*. IGI Global, United States, pp. 73-95, 2023.

[19] P. Suetens. "Medical image analysis". In: *Fundamentals of Medical Imaging*. Cambridge University Press, United Kingdom, 2009.

[20] R. Nock and F. Nielsen. "Statistical region merging". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, no. 11, pp. 1452-1458, 2004.

ORIGINAL RESEARCH ARTICLE UHD JOURNAL OF SCIENCE AND TECHNOLOGY

Enhanced Kidney Stone Detection and Classification Using SVM and LBP Features

Hawkar K. Hama[1], Hamsa D. Majeed[2], Goran Saman Nariman[3]

[1]Department of Computer Science, College of Basic Education, University of Sulaimani, Kurdistan region, Iraq, [2,3]Department of Information Technology, College of Science and Technology, University of Human Development, Kurdistan Region, Iraq

ABSTRACT

Nephrolithiasis is a scientific term that refers to kidney stones and means the formation of crystal concretions in the kidney. It is considered a widespread situation that affects millions of people worldwide. Those stones can cause serious discomfort to infected people, especially when they traverse the urinary system, although, the big stones may need a surgical intervention. Various systems are already in use to address kidney stones, including ultrasound imaging for detection, extracorporeal shock wave lithotripsy (ESWL) for non-invasive stone fragmentation, and ureteroscopy for surgical removal, showcasing the advances in medical technology for managing this condition. This study presents an approach for detecting stones in the affected kidney. A public dataset has been employed in this work, containing (2370) images of healthy and affected kidneys. The dataset was utilized to train the proposed approach for the aim of stone detection. To achieve high detection accuracy, we implemented two key phases before classification. The preprocessing phase enhances image quality by reducing noise using a median filter and improving contrast through contrast stretching and tone enhancement. The segmentation phase follows, accurately identifying the kidney's edges and regions of interest for effective feature extraction. The Local Binary Pattern (LBP) technique, combined with the support vector machine (SVM) algorithm serves as the primary components of the proposed model. The feature extraction comes into action through the LBP technique as a preparation step for the SVM classifier to complete the stone detection process. The approach introduced in this paper has the potential to enhance detection accuracy and efficiency. Furthermore, it could be used as an early detection tool to identify potential cases, thereby helping to prevent complications and adverse outcomes. This method aims to improve on the traditional manual process employed by radiologists, which could be described as time and effort consumption rather than the exposure of the interpretations. The obtained results were compared with the most relevant approaches in the field of kidney stone detection, demonstrating the model's effectiveness in achieving the desired goal with a diagnostic accuracy of 96.37% for kidney stones.

Index Terms: Medical Image Analysis, CT Images, Local Binary Pattern, Support Vector Machine, Kidney Stones

1. INTRODUCTION

Image processing techniques have been proven as powerful tools in the field of medical image analysis due to their efficiency in improving the quality of those images along with the ability to extract useful information from them. Furthermore, they have been combined with machine-learning techniques; and together they achieve a quite noticeable progress in different domains, precisely in healthcare applications. Since kidney stones are considered a serious threat to people's health, many developments and research applied in this field to reduce that threat to the minimum, and over time early detection becomes a necessity for the diagnosis. The researchers invested a lot of their time and thoughts for that matter through many approaches that were based on image processing, machine learning, or both, including the segmentation and feature extraction techniques [1], [2].

Access this article online	
DOI: 10.21928/uhdjst.v9n1y2025.pp10-17	**E-ISSN:** 2521-4217 **P-ISSN:** 2521-4209

Corresponding author's e-mail: Hawkar K. Hama, Department of Computer Science, College of Basic Education, University of Sulaimani, Kurdistan region, Iraq. E-mail: hawkar.hama@univsul.edu.iq

Received: 15-08-2024 Accepted: 28-12-2024 Published: 13-01-2025

More advancements have been established by merging a variety of those techniques to support early diagnosis accurately. One of the powerful algorithms that achieve such improvement is the Support Vector Machine (SVM) through CT images. SVMs are valued for their ability to create clear decision boundaries, making them effective for binary and multiclass classification tasks. Its capability of classifying the kidney stones correctly was obvious and greatly participated in patient care [3], [4].

The spread of the problem of kidney stones has become the focus of attention of many researchers. The worldwide research community has recognized the urgent need for accurate and effective techniques for the rapid detection of kidney stones to reduce the potential for complications and patient uneasiness.

To overcome the penalties of the usual diagnosis procedure through the established methods by the radiologists which take a quit amount of time and effort with the potential of interpretations and complications that could lead to a negative affection on the patient, the suggested model can offer a precise improvement in early kidney stones detection and prevent any expected inconvenience.

The early detection and the accurate diagnosis of kidney stones are the main objectives of this work. A suggested approach presented in this study for that purpose utilizes a Support Vector Machine (SVM) algorithm with the Local Binary Pattern (LBP) technique for the premature and precise detection of kidney stones avoiding unnecessary consequences and providing proper medical care. The model was implemented using MATLAB (R2021a) starting from fetching the input images through a series of preprocessing for the image's quality enhancement and crossing the segmentation phase for Region of Interest (ROI) determination. Fifty-nine features extracted by LBP to be fed into SVM for classification. The upcoming sections present a literature review to provide background and context, the proposed method detailing our approach, implementation, and results to demonstrate our findings, and a conclusion summarizing key insights and future directions.

2. LITERATURE REVIEW

Introduced in this section is a diversity of remarkable approaches and contributions including the specifics of the proposed models in their works with the extracted phases that contributed to the enhancement of the stone's recognition accuracy.

The frequent appearance of stone issues in patient's kidneys motivated the authors in [5] to employ Computer-Aided Detection (CAD) algorithms for stone detection using CT images. Their approach highlighted the problem of organ determination, directing attention to the importance of segmenting the image and emphasizing the region of interest (ROI) and how this addition will have an effect on the stone detection accuracy. Ultrasound images are employed by [6] for stone detection, for that they added a preprocessing phase to the used images along with a Median filter for image quality enhancement and noise removal, the image was then segmented using a morphological segmentation, the confusion matrix gathered by the Gray Level Co-occurrence Matrix (GLCM) used then for model assessment with a rate of 90%. The authors in [7] also utilized ultrasound images, incorporating a preprocessing phase into their model. In addition to applying morphological procedures, they enhanced the approach by integrating fuzzy masking. Then entropy-based segmentation is used to define the (ROI). For final classification, the SVM and KNN algorithms were used, and the classification accuracy results were 89% and 84% for the KNN and SVM classifiers, respectively.

The author in [8] distinguished between healthy tissues and those surrounding abscesses, fibrosis, and stones using a watershed segmentation technique, and as a tissue index LBP was used. To describe the shape and wrongdoing of kidney conditions, such as the compositional spectrum was used as statistical features, in addition to the geometric features. The accuracy of the system was evaluated as follows: 88.4% for LBP and 91.36% for the synthetic spectrum, while the combination of the local binary spectrum and the synthetic spectrum achieved 95%.

To enhance diagnosis and treatment, a study conducted in [9], [10] focused on improving the identification of kidney stone types. The authors carried out a pilot study to explore the classification of kidney stones using *in vivo* images obtained during ureteroscopy procedures. These images were analyzed, and visual features used by urologists to differentiate stone types were encoded into vectors for kidney stone surface and cross-section. The feature vectors combined the color features and the texture features where LBP was employed for extraction. The classification was performed in [9] using Random Forest and ensemble K Nearest Neighbor classifiers, achieving 89% overall accuracy. At the same time, in [10] They compared the performance of six shallow machine-learning methods and three deep-learning architectures, SVM was used as one of the methods with precision results of 79% for the mixed kidney stone patches (surface and section).

In addition, CT images of a kidney with a stone were examined by [11] through a model consisting of neural network-based and SVM-based classification, while preprocessing techniques were employed for noise removal and GLCM for feature extraction, for the ANN-based model achieved 85%, while using SVM gave an accuracy of 95% with fewer features and 99% with full features.

The combination between neural network and SVM continued in [12] where the authors proposed a model named a "Hybrid Deep VGN-19 and Binary SVM (HDVS)," deep learning techniques were employed for extracting the features whereas the SVM served as a classifier, a metric of precision, F1 score, recall and accuracy used for the system evaluation which achieved a rate of 99.89% for accuracy.

A renal calculi identification explored by [13] by presenting an SVM-based model through a small number of only 250 ultrasound images, median filter employed to reduce the image's noise, K-means clustering and GLCM were used for segmentation and extracting features, respectively, the model attained an accuracy of 98.8%.

Another neural-svm combination proposed by [14] where X-Ray images were examined using multiple machine learning methods such as Random Forest, K-Nearest Neighbor, Decision Trees, Multilayer Perceptron, CNN, and Naive Bayes (BernoulliNB) besides SVM, the approach assessed through F1 score, precision, and recall. The outcomes of SVM were 92.4%, 85.8%, and 85.8%, respectively.

In another study, an approach presented by [15] for kidney disease identification, the approach used SVM and ANN as classifiers to compare the performance of the two algorithms based on the accuracy and time consumption, the resultant outcomes revealed that ANN was more accurate in detection with 87.70% accuracy than SVM which achieved 76.32%. By leveraging preprocessing followed by a combination of LBP for feature extraction and SVM for classification, our approach demonstrates better accuracy and efficiency, particularly in handling noisy images, as reflected in the results. This advantage addresses the limitations of the compared methods, which often lack feature representation and adaptability, leading to lower accuracy.

3. THE PROPOSED METHOD

The proposed method in this study leverages supervised learning techniques for accuracy enhancement regarding kidney stone detection in CT scan images, a critical aspect of urolithiasis diagnosis.

The proposed architecture of the model in this work as presented in Fig. 1 has been established through four primary stages. As mentioned in [16], the whole possible implementable recognition process consists of four consecutive phases, namely: Pre-processing (P), Segmentation (S), Feature Extraction (F), and Classification (C). The kidney stone detection process in this work is depicted as (PSFC) indicating the employment of the phases.

The implementation was carried out using MATLAB (R2021a), the process starts with image preprocessing to enhance the contrast and remove noise, followed by segmentation to detect the objects and their boundaries in the images. Then, feature extraction is applied using the LBP technique. The extracted features were used afterward for SVM training, which was used to predict the presence of kidney stones in new CT images. This section provides a step-by-step illustration and a visual diagram for replicating the implementation.

3.1. Input CT-Image

This study utilized a public dataset [17] of CT images, with 5077 images of normal kidneys and 1,377 of kidneys with stones. The data set contains other images related to kidney diseases that are not included in this study, so they are discarded for discussion. The total number of images used in this work is 6454 from the dataset. The images are in JPG format and have a dimension of 512 × 512 pixels.

As mentioned, the number of images in these two classes is not equal. To address the challenge of this imbalanced dataset in SVM, there are several techniques, such as resampling and class weights techniques. Resampling techniques include oversampling and undersampling. Oversampling involves generating additional samples (images) of the minority class to rebalance the dataset while undersampling reduces the

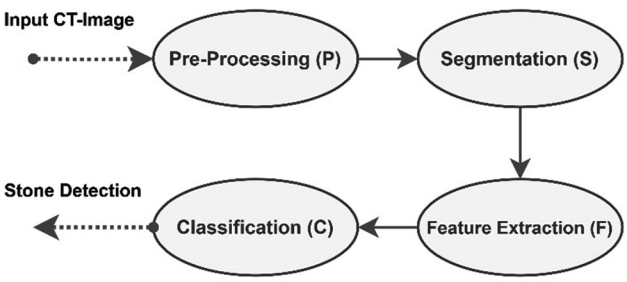


Fig. 1. The proposed model architecture.

number of samples (images) in the majority class. These techniques aim to create a more equitable distribution of classes for SVM training. However, while resampling can be effective, it may lead to information loss or overfitting. The class weight technique assigns varying levels of significance to each class during SVM training, with greater weights assigned to the minority class and lesser weights to the majority class, this means that the SVM prioritizes the correct classification of the minority class (the stone class) without altering the dataset's composition.

In this study, the class weight technique has been adopted for its distinct advantages over resampling methods, including:

- Efficiency: Class weights are computationally more efficient compared to resampling techniques. Resampling involves creating duplicate samples, which can significantly increase the dataset size, leading to longer training times. In contrast, class weights only affect the loss function during training without changing the dataset size.
- Preservation of Data: Resampling methods such as oversampling and undersampling can result in loss of information or potential overfitting, as they duplicate or remove data points. Class weights do not alter the data distribution, ensuring that all original data points are considered during training.
- No Need for Data Generation: Resampling may require generating synthetic data points or discarding data, which can introduce bias or inaccuracies. Class weights do not require creating synthetic data, reducing the risk of introducing artifacts.

For a more comprehensive view of the dataset distribution, Table 1 provides a detailed breakdown of the dataset distribution, including the total number of images in each class, as well as their division into training and testing subsets. In addition, the dataset was partitioned into two distinct subsets: The training set, encompassing 80% of the images, and the test set, encompassing the remaining 20%. These images served as input for various computer vision techniques to extract relevant information and features.

3.2. Pre-processing

Numerous studies have substantiated that the model's accuracy rate is inherently dependent on the output quality of the preprocessing phase. This study seeks to further corroborate this notion by introducing a preprocessing phase comprising three stages. These stages are deployed to enhance the quality of the input image, consequently leading to an improvement in the accuracy of kidney stone identification. The pre-processing steps are delineated as follows:

3.2.1. Median filter

In this study, the median filter was applied during preprocessing to reduce noise in the CT images while preserving critical edge details. This step ensured that the kidney's boundaries and texture features were retained, improving the accuracy of subsequent segmentation and feature extraction phases.

3.2.2. Contrast enhancement

Contrast enhancement was utilized to improve the visibility of kidney structures and potential stones in the CT images. This process involved two steps: Contrast stretching to distribute brightness uniformly across the dynamic range, followed by tone adjustment to highlight key areas such as shadows, mid-tones, and highlights. These enhancements made critical features more distinguishable, aiding in accurate edge detection, and feature extraction.

3.3. Segmentation

Segmentation is frequently used to distinguish between objects and boundaries in images, it is considered an essential component in medical image analysis, it makes medical data representation easier to understand and facilitates the diagnosis of a variety of disorders.

To identify kidney stones and clarify their boundaries, Canny edge detection was used in this study. One of the characteristics of this technology is that it can separate important visual parts from those surrounding them architecturally and in a largely unified manner. Thus, it helps in reducing the amount of data that must be processed. This enables it to detect the edges of objects to be focused on, which enhances the success requirements for the system that uses its features.

3.4. Feature Extraction

The feature extraction process is considered one of the main phases in the used system. This phase focuses on transforming segmented kidney images into numerical

TABLE 1: Dataset usage of CT kidney images

Type of image	No. of images	Training Set (80%)	Testing Set (20%)
Normal kidney	5077	4062	1015
Kidney with stone	1377	1102	275
Total	6454	5164	1290

representations, extracting useful information from the region of study, and building it out into statistics in a more comprehensible formation. This study employed the technique of LBP for the feature extraction phase. The technique was first presented in 1996 [18], and since it has been widely applied in texture analysis across various fields, image analysis was one of them, the technique analyzes the grayscale value of each pixel in the image's region of interest, comparing it to neighboring pixel values within a defined radius. It generates a binary pattern for each pixel based on these comparisons, which is then converted into a decimal value that describes the local texture.

In this study, LBP was applied to segmented CT images obtained after pre-processing and edge detection to capture textural variations in the kidney regions indicative of stones, the technique has a valuable achievement and approved its worth in local texture analysis.

A total of 59 features were extracted from each segmented image, balancing sufficient detail with computational efficiency. These features captured local texture differences crucial for distinguishing between normal kidneys and those with stones and were subsequently used by the SVM for classification.

This systematic feature extraction ensures that the most relevant information for kidney stone detection is captured, enabling accurate and reliable classification results.

3.5. Support Vector Machine

In the supervised machine learning field of study, the support vector machine (SVM) algorithm shines as an effective classifier. Its mechanism is based on separating data points into classes in a hyperplane through a high-dimensional feature space. The basic working mechanism was presented by Vapnik in 1998 [19], and he is considered to have laid the foundation for the general theory of statistical learning. Since then, the theory has witnessed many expansions and developments while maintaining the basic goal, which is to determine the excess level to increase the margin between different classes. In this work, SVM was employed to identify kidney stones from CT images. The algorithm was trained on a public dataset of labeled CT images using MATLAB's built-in functions and toolboxes for image processing and machine learning.

4. IMPLEMENTATION AND RESULTS

Experiments have been carried out to evaluate the performance and applicability of the proposed model implemented using MATLAB. The proposed model was performed on two sets of images from a public dataset of kidney's CT images, normal kidney set and kidney with stone set.

A comprehensive workflow diagram, as depicted in Fig. 2, outlines the implementation of our proposed method. This workflow provides a visual representation of the entire process, allowing a clear understanding of the methodology.

The images underwent a sequence of pre-processing procedures aimed at enhancing the input image's quality, ultimately resulting in heightened accuracy for kidney stone identification. These pre-processing steps encompassed the conversion of the images to grayscale, preserving a spectrum of gray shades. Subsequently, a median filter was applied to reduce noise, and contrast enhancements were employed to enhance the visibility of objects within the image. Fig. 3 illustrates the results of this phase, particularly for a kidney with a stone.

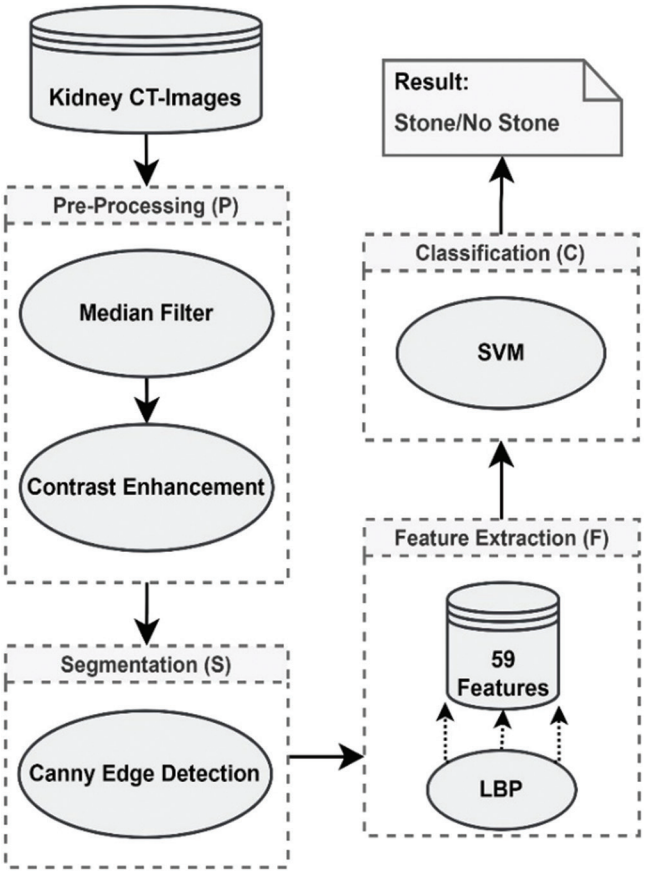


Fig. 2. Overview of the proposed model with applied phases and techniques.

Following the pre-processing stage, the segmentation phase was executed to identify objects and boundaries within the images. The outcome of this segmentation process is depicted in Fig. 4.

After applying Canny edge detection to segment a kidney CT image, the resulting image clearly defines and outlines the kidney structure. This high-contrast image significantly aids in the accurate identification of kidney stones.

After segmentation, comes the next necessary stage, which is feature extraction. The LBP technique was used for this purpose because it is an effective way to evaluate textures and patterns in images. When it comes to identifying kidney stones, it shows how important this technique is to identify important features and support the classification process. LBP focuses on storing local texture information to extract useful features from previously segmented kidney images that are necessary for the process of distinguishing between stone-affected kidneys and healthy kidneys. Choosing a feature extraction technique is an important stage when building the system because it has a direct impact on the success and accuracy of the classification process.

After extracting important features using the LBP technique, the kidney images are now transformed into a set of selective features that can be submitted to the classification stage, represented by the SVM technique, which has proven effective in the field of machine learning and is known for classifying data into separate groups according to the extracted feature vectors. The relevant features extracted by LBF are fed to the classification phase where SVM is applied as a classifier for training which yields images to be classified into "with stone" or "without stone" kidneys. The two technologies were used together to ensure a high level of accuracy, and this was proven, as the results indicated the detection of kidney stones with an accuracy of 96.37%, which is a new indicator that indicates the system's potential and ability to differentiate between a healthy kidney and those with stones efficiently. The classification outcomes are visualized using a pie chart as shown in Fig. 5, this chart illustrates the classification results for each class based on the proposed approach.

These results highlight the ability of the machine learning model to predict health outcomes and the SVM model to detect kidney stones. Further studies may focus on enhancing and utilizing these models in different contexts.

Furthermore, a comparative analysis was carried out to assess the accuracy performance of our approach concerning earlier research in this area. Table 2 provides a comprehensive comparison of the classification accuracy achieved by the proposed method alongside previous works. The accuracy values reflect the effectiveness of each approach in distinguishing between normal kidney images and those containing stones. The "Dataset" column in Table 2 categorizes the datasets based on their origin as presented by [16], distinguishing between publicly available datasets

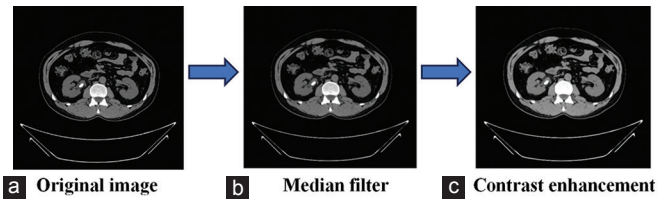


Fig. 3. (a-c) The outcomes of the pre-processing phase.

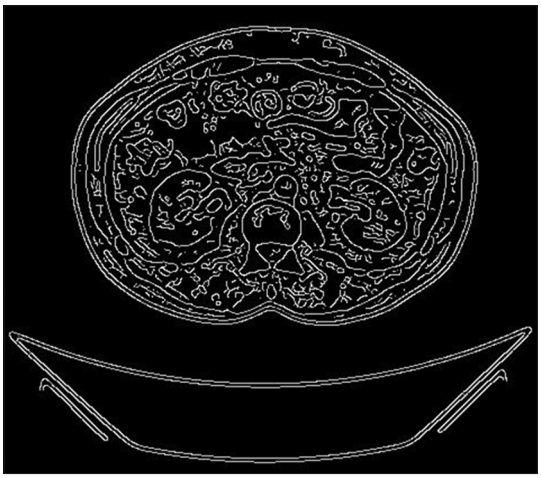


Fig. 4. Segmented image.

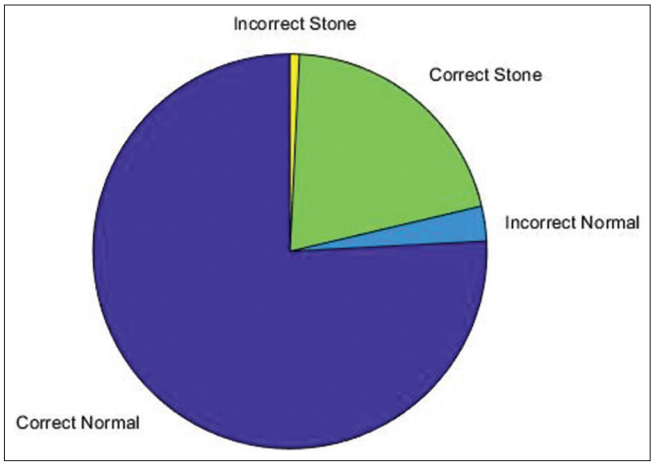


Fig. 5. Classification results for each class based on the proposed model.

TABLE 2: Comparison of classification accuracy

Approach	Dataset	FE Technique	Classifier	Accuracy (%)
[7]	Public	PCA	SVM	84
[8]	[17]	LBP	-	88.4
[9]	Self-constructed	LBP	Random Forest Ensemble KNN	89
[10]	Self-constructed	LBP	SVM	79
The proposed method	[17]	LBP	SVM	96.31

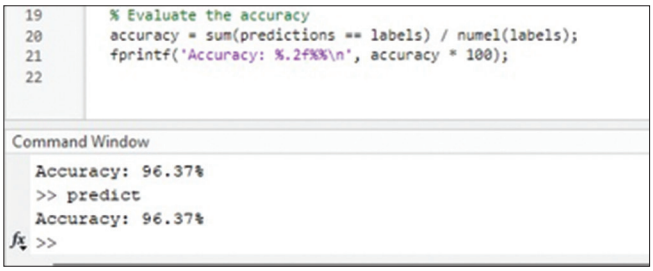


Fig. 6. Screenshot of the simulation reveals the result.

and those that are self-constructed, the datasets are classified whether they are publicly available or constructed by authors specifically for their approach. It is worth noting that some of the approaches listed in the table utilized the same public dataset that we employed in our method, as indicated by the corresponding references.

Table 2 displays a comparison of classification accuracy among different approaches for kidney stone detection. It provides information about the dataset used (whether public or self-constructed), the feature extraction technique employed, the classifier used, and the accuracy achieved for stone detection in each approach.

All the approaches are in common with our proposal either with the used public dataset, the feature extraction technique, or the classifier. Among these, as shown in Fig. 6 the proposed method demonstrates the highest accuracy of 96.31% among all the listed approaches, signifying its effectiveness in accurately detecting kidney stones.

5. CONCLUSION

In this study, we aimed to tackle the critical challenge of kidney stone detection. Our proposed approach offers an efficient solution for the rapid and accurate identification of kidney stones, which can be a root cause of various health issues. Our investigation revealed the effectiveness of the SVM model in predicting the presence of stones, highlighting its potential to guide preventive and responsive healthcare strategies. In addition, we developed a machine learning model that harnesses the power of an SVM and utilizes LBP feature extraction to enhance accuracy by effectively extracting 59 features from segmented CT images to capture local texture details which are essential for classification. To achieve the desired result, the main process followed a systematic sequence: Pre-processing for noise reduction and contrast enhancement, segmentation to identify the kidney's boundaries and ROI, feature extraction using LBP, and classification using SVM. The model was trained and validated using the publicly available CT KIDNEY DATASET: Normal-Cyst-Tumor and Stone ensuring replicability and standardization. This model was deployed for stone prediction, yielding an accuracy of 96.31%. In addition, the results were compared with related approaches, where this approach outperformed existing methods in terms of accuracy. The outcomes of this research hold substantial promise for public health, particularly in the realm of kidney stone detection.

REFERENCES

[1] S. Singh, H. Singh, G. Bueno, O. Deniz, S. Singh, H. Monga and P. N. Hrisheekesha. "A review of image fusion: Methods, applications and performance metrics". *Digital Signal Processing*, vol. 137, pp. 104020-104020, 2023.

[2] Y. M. Y. Abdallah and T. Alqahtani. "Research in medical imaging using image processing techniques". In: *Medical Imaging-Principles and Applications*. BoD-Books on Demand, Norderstedt, 2019.

[3] M. Javaid, A. Haleem, R. P. Singh, R. Suman and S. Rab. "Significance of machine learning in healthcare: Features, pillars and applications". *International Journal of Intelligent Networks*, vol. 3, pp. 58-73, 2022

[4] A. Soni and A. Rai. "*Kidney Stone Recognition and Extraction Using Directional Emboss and SVM from Computed Tomography Images*". Institute of Electrical and Electronics Engineers, United States, 2020.

[5] V. Nazmdeh and S. S. Esmaili. "A review of methods for detection and segmentation of kidney stones from CT scan images using image processing method". *International Journal of Cybernetics and Cyber-Physical Systems*, vol. 1, no. 2, p. 157, 2022.

[6] R. Goel and A. Jain. "Improved detection of kidney stone in ultrasound images using segmentation techniques". In: *Advances in Data and Information Sciences*. Springer Nature, Germany, pp. 623-641, 2020.

[7] J. Verma, M. Nath, P. Tripathi and K. K. Saini. "Analysis and identification of kidney stone using Kth nearest neighbour (KNN)

and support vector machine (SVM) classification techniques". In: *Pattern Recognition and Image Analysis*. Vol. 27. Springer, Germany, pp. 574-580, 2017.

[8] A. H. Ali. "Studying the kidney textural using statistical features and local binary pattern". *Journal of Al-Nahrain University Science*, vol. 20, no. 4, pp. 64-76

[9] A. Martinez, D. H. Trinh, J. El Beze, J. Hubert, P. Eschwege, V. Estrade, L. Aguilar, C. Daul and G. Ochoa. "*Towards an Automated Classification Method for Ureteroscopic Kidney Stone Images using Ensemble Learning*". HAL (Le Centre pour la Communication Scientifique Directe), IEEE, United States, 2020.

[10] G. Ochoa-Ruiz. "*On the in vivo Recognition of Kidney Stones Using Machine Learning*". Cornell University, United States, 2022.

[11] P. Chak, P. Navadiya, B. Parikh and K. C. Pathak. "Neural network and SVM based kidney stone based medical image classification". In: *Communications in Computer and Information Science*. Springer, Germany, pp. 158-173, 2020.

[12] K. Somasundaram, S. M. Animekalai and P. Sivakumar. "*An Efficient Detection of Kidney Stone Based on HDVS Deep Learning Approach*". ICCAP, Pennsylvania, 2021.

[13] S. Selvarani and P. Rajendran. "Detection of renal calculi in ultrasound image using meta-heuristic support vector machine". *Journal of Medical Systems*, vol. 43, no. 9, p. 300, 2019.

[14] I. Aksakalli, S. Kaçdioğlu and Y. S. Hanay. "Kidney X-ray images classification using machine learning and deep learning methods". *Balkan Journal of Electrical and Computer Engineering*, vol. 9, pp. 144-151, 2021.

[15] P. Sinha and P. Sinha. "Comparative study of chronic kidney disease prediction using KNN and SVM". *International Journal of Engineering Research*, vol. V4, no. 12, 2015, p. 608.

[16] H. D. Majeed and G. S. Nariman. "Construction of alphabetic character recognition systems: A review". *UHD Journal of Science and Technology*, vol. 7, no. 1, pp. 32-42, 2023.

[17] "*CT Kidney Dataset: Normal-Cyst-Tumor and Stone".* Available from: https://www.kaggle.com/datasets/nazmul0087/ct-kidney-dataset-normal-cyst-tumor-and-stone [Last accessed on 2025 Jan 08].

[18] T. Ojala, M. Pietikäinen and D. Harwood. "A comparative study of texture measures with classification based on featured distributions". *Pattern Recognition*, vol. 29, no. 1, pp. 51-59, 1996.

[19] V. N. Vapnik. "*Statistical Learning Theory*". Wiley, Cop, New York, 1998.

REVIEW PAPER | UHD JOURNAL OF SCIENCE AND TECHNOLOGY

Exploring Post-Quantum Cryptography: Evaluating Algorithm Resilience against Global Quantum Threats

Tara Nawzad Ahmad Al Attar, Mohammed Anwar Mohammed, Rebaz Nawzad Mohammed

Department of Computer Science, College of Science, University of Sulaimani, Sulaymaniyah, Iraq

ABSTRACT

Cryptographic algorithms perform a vital part in protecting information in general and safeguarding digital platforms. Nevertheless, improvements in quantum computing pose important concerns to traditional cryptographic approaches, demanding the development of quantum-resistant explanations. This study offers an inclusive investigation of post-quantum cryptographic algorithms, assessing their flexibility, competence, and practicality in justifying quantum risks. Through an equivalent approach, the research identifies optimistic applicants for upcoming cryptographic standards. Moreover, the study highlights the international essential for embracing these algorithms to ensure secure communication and data protection in the quantum era. These conclusions aim to notify the progress of strong cryptographic systems that address the appearing objections of quantum technologies.

Index Terms: Post-Quantum Cryptography, Quantum Threats, Spatial Implications, Cryptographic Algorithms

1. INTRODUCTION

As technology endures to advance, important concerns arise concerning the security of sensitive information and the veracity of digital communications across numerous platforms. A fundamental tool for completing data disclosure and integrity is the use of cryptographic algorithms. These algorithms provide robust protection against unauthorized access, ensuring that data remains secure even when intercepted during transmission or storage [1]. However, as technology develops, new computational threats to traditional cryptographic systems also emerge. Quantum computing, as a novel computational model, possesses the potential to compromise many widely used encryption methods, rendering current cryptographic techniques vulnerable [2]. Nonetheless, Quantum threats are existing and they state the potential dangers posed by quantum computers to traditional cryptographic systems and data security. Quantum computers, exploiting principles of quantum mechanics, can resolve specific computational difficulties exponentially faster than classical computers, rendering many existing cryptographic methods vulnerable [3]. In response to this challenge, post-quantum cryptography (PQC) states to a division of cryptography that emphasizes on emerging cryptographic algorithms that are resilient to attacks by quantum computers. Unlike traditional cryptographic methods that rely on mathematical problems resolvable by quantum algorithms (e.g., Shor's algorithm for factoring), it aims to protect communications and data against these advanced computational threats [4]. PQC also been developed to safeguard against the unique threats posed by quantum computing. This branch of cryptography focuses on algorithms that are designed to resist attacks from advanced quantum systems [5]. It is important to recognize that PQC serves as a crucial line of defense against the

Access this article online

DOI: 10.21928/uhdjst.v9n1y2025.pp18-28 | **E-ISSN:** 2521-4217 | **P-ISSN:** 2521-4209

Corresponding author's e-mail: tara.ahmad@univsul.edu.iq

Received: 27-11-2024 Accepted: 11-01-2025 Published: 25-01-2025

capabilities of quantum computers, which can undermine the security of classical cryptographic methods due to their immense computational power [6]. Joining Geographical Information System (GIS) into the investigation of PQC and quantum threats enhances a vital spatial viewpoint, allowing participants to rank and execute secure systems adequately. It provides functional understandings to policymakers, security specialists, and organizations intending to defend digital infrastructures internationally [7]. However, Spatial significance using (GIS) in the context of PQC and quantum threats include analyzing, visualizing, and understanding the geographic allocation and influence of quantum computing progressions, weaknesses, and the acceptance of quantum-resistant cryptographic systems worldwide [8].

This paper presents an in-depth analysis of post-quantum cryptographic algorithms, investigating their strength against quantum raid [9], It concentrates on the theoretical and practical suggestion of these algorithms, evaluating their capability to safeguarding digital communication in the face of developing quantum risks. By rating the durability and deficiencies of numerous algorithms, the paper aims to direct the ragged placement of cryptographic explanations that safeguard robust security in a post-quantum world [5]. These understandings intend to support international digital security policies and notify decision makers for the implementation of PQC [10].

2. LITERATURE REVIEW

2.1. Evolution of Cryptography and Its Challenges

Cryptography has developed considerably over the centuries, progressing from basic ciphers, such as Caesar's cipher to the detailed cryptographic procedures that are used today. Primary cryptographic approaches are primarily intensive on safeguarding communication, such as military messages during wartime [11]. Throughout time, cryptography progressed with performances, such as the Vigenère cipher and the Enigma machine, which pioneering more complexity and contribution to cryptographic developments [12]. The digital era introduced public-key cryptography in the 1970s, this uprising development by Diffie and Hellman abolished the need for pre-shared secret keys, allowing safe communication over open channels [13]. These historic developments placed the foundation for current cryptographic applications, but as technology advanced, so did the difficulty and possible risks to digital systems.

2.2. The Quantum Computing Challenge

The escalation of quantum computing offers a novel challenge to the protection of traditional cryptographic systems, especially public-key algorithms, such as RSA and Diffie-Hellmann [14]. Quantum computers are able to solve mathematical difficulties exponentially faster than standard computers, which poses a significant risk to widely used encryption methods [15]. Shor's algorithm, a quantum algorithm skilled for factoring large numbers and solving the discrete logarithm complication, hovers to concentrate on traditional cryptographic systems vulnerabilities [14]. In addition, quantum algorithms, such as Grover's algorithm compromise balanced cryptographic systems such as AES by reducing their actual security strength, this has urged an ambition to emerging cryptographic approaches that can endure the power of quantum computing [16].

2.3. PQC

PQC refers to cryptographic algorithms marked to persist security against the computational ability of quantum computers. PQC intends to address the boundaries of traditional cryptographic structures by depending on mathematical complications, that are resilient to quantum attacks [17]. Experts have discovered numerous mathematical structures for PQC, including lattice-based cryptography, code-based cryptography, multivariate polynomial cryptography, and hash-based cryptography. The literature on PQC has expanded significantly in the past decade, as experts attempt to develop new cryptographic standards that can replace vulnerable classical algorithms [18].

2.4. Lattice-Based Cryptography

Lattice-based cryptography has garnered important consideration due to its robust resistance to quantum attacks. The security of lattice-based structures depending on the solidity of complications, however Notable examples of these hard problems are Shortest Vector Problem (SVP), and Learning With Errors (LWE), which are supposed to be statistically difficult in quantum computers [19]. Lattice-based cryptography has the plus of comparatively efficient encryption and decryption processes, though it faces challenges in slower key generation times and larger key sizes.

2.5. Code-Based Cryptography

Code-based cryptography is another remarkable competitor for post-quantum security. This technique uses error-correcting codes, and its security is based on the difficulty of decoding random linear codes, a problem that remains hard for quantum algorithms to solve [20], despite the fact that code-based cryptography suggests robust resistance to quantum risks and attack, but the key sizes

are usually large, which can present challenges for applied implementation [21]. Despite this, code-based cryptographic schemes are considered highly capable for quantum-resilient encryption.

2.6. Multivariate Polynomial Cryptography

Multivariate polynomial cryptography counts on algebraic equations, mainly multivariate polynomials, to generate encryption schemes. These schemes are measured to be secure against quantum algorithms, as solution systems of multivariate polynomials are computationally difficult to solve in quantum computers [22]. Nevertheless, the major disadvantages of multivariate polynomial cryptography are its fairly slow encryption and key generation processes, which can delay the practical placement in performance-sensitive applications.

2.7. Hash-Based Cryptography

Hash-based cryptography offers a distinct method, using cryptographic hash functions as the groundwork for security. The main benefit of hash-based schemes is their resistance to quantum strikes, as the problem of shifting cryptographic hash remains problematic even for quantum computers [23]. Nevertheless, the hash-based schemes are frequently inadequate by greater signature extents and slower signing times, which reduce their applied competence [24].

2.8. Comparison of Post-Quantum Cryptographic Approaches

Each of the above-mentioned post-quantum cryptographic methods has its strengths and weaknesses, and an important form of this literature has focused on comparison between these methods to detect the most optimistic candidates for upcoming cryptographic standards. Lattice-based cryptography, for example, suggests a good stability of security and efficiency, though it is often disapproved for slower key generation. Code-based cryptography is extremely secure but undergo from large key sizes that make it impractical in some applications. Multivariate polynomial cryptography offers strong security but struggles with slow performance, while hash-based schemes are resilient to quantum attacks but are limited by large signatures and slow signing processes [25].

The challenge for the cryptographic group lies in choosing or developing the best set of PQC algorithms that balance security, efficiency, and practical applicability. Researchers continue to discover cross approaches, such as combination essentials from lattice-based and code-based cryptography, to overcome the boundaries of individual methods.

2.9. Geographic Information Systems (GIS) and Quantum-Resilient Cryptography

In addition to algorithmic developments, GIS play a vital part in recognizing regional weaknesses and supporting the international placement of post-quantum cryptographic answers. GIS can offer spatial analysis to identify areas at larger risk of quantum attacks and guide the application of targeted cryptographic solutions in these regions [26]. By joining GIS with post-quantum cryptographic systems, researchers and policymakers can ensure that the security of both localized and international systems is enhanced as the quantum computing era approaches [27]. Eventually, the development of quantum computing has piloted in new encounters for cryptography, requiring the development of post-quantum cryptographic systems. Lattice-based, code-based, multivariate polynomial, and hash-based cryptography represent the forefront of this study, each with separate benefits and trade-offs. As quantum computing improvements, ongoing researches are essential to identify, refine, and standardize post-quantum cryptographic approaches. GIS also provides an important tool in supporting the international execution of these resolutions, safeguarding that cryptographic systems are resilient to quantum attacks across diverse regions. Constant innovation and association will be vital in safeguarding digital communication in the quantum computing era.

3. METHODOLOGY

3.1. Research Design

This study intends to estimate the theoretical resilience of post-quantum cryptographic algorithms in contrast of possible quantum computing risks. It also highlights a wide range of theoretical investigations, shared with a serious evaluation of current literature, to measure the cryptographic ethics, efficiency, and applicability of these algorithms. The main attention is on understanding their potential in real-world distribution and their inferences for cybersecurity in the quantum era. The study will address the following key components:

- Scope: The study explores a nominated set of post-quantum cryptographic algorithms, including lattice-based, hash-based, and code-based cryptographic methods. The valuation is based on theoretical perceptions resulting from academic literature and technical reports. This work does not include experimental testing or geopolitical analysis, instead, it is relying on secondary data to draw conclusions about the algorithms' robustness and practicality.

- Objectives: The main aim is to measure the resilience of post-quantum cryptographic algorithms against possible quantum attacks. This will include evaluating both the strengths and weaknesses of these algorithms when exposed to quantum-based cryptography, while also identifying their suitability for real-world deployment. In addition, the study aims to provide actionable insights for cryptographical academics, experts, and officials, apprising future cryptographic standards and policies in the expectancy of quantum-era fears.

The key objectives are:
- To analyze the theoretical foundations of post-quantum cryptographic algorithms.
- To evaluate their strengths and limitations in the context of quantum computing advancements.
- To provide insights for researchers and policymakers in the field of PQC.
- Target audience: The verdicts of this study will be of specific interest to cryptographic researchers, cybersecurity experts, and policy consultants to elaborate in the development of next-generation cryptographic systems. The study will also oblige to enlighten international negotiations on the combination of PQC into international security.

3.2. Analytical Approach

This research will employ a cross-methodological framework that blends theoretical analysis with empirical testing:
1. Theoretical analysis: A detailed valuation of the mathematical values and cryptographic constructions behind nominated post-quantum algorithms. This analysis draws on peer-reviewed literature, industry whitepapers, and technical standards documentation.
2. Comparative assessment: Algorithms are compared based on the system of measurement (key metrics) such as theoretical security, computational productivity, and practicality for real-world applications. This comparative approach enables the identification of trade-offs and optimal solutions for different cryptographic needs.

3.3. Data Collection

This study will collect data from various confident sources to confirm a comprehensive analysis of post-quantum cryptographic algorithms and their resilience to quantum threats. Key data sources will include:
1. Academic literature: Peer-reviewed articles and conference papers will provide foundational insights into the theoretical resilience of the algorithms.
2. Whitepapers and technical documents: Industry and standards organization whitepapers offer applied viewpoints on algorithm implementation and efficiency.
3. Industry and government reports: from cybersecurity organizations and government bodies will offer context on quantum computing advancements and projected quantum threats.

Together, these data sources will create an accomplished groundwork for measuring post-quantum cryptographic algorithms, allowing a fine understanding of their flexibility in a world where quantum fears are increasingly possible.

3.4. Quantitative and Qualitative Analysis

- Metrics selection: Key metrics, such as security strength, computational complication, and applied application, will be nominated to assess each algorithm's resilience to quantum attacks.
- Quantitative analysis: Mathematical representations and simulations will be applied to quantitatively measure algorithm performance. This contains hard computational and mathematical examinations to measure algorithmic efficiency and resistance to quantum-based threats.
- Qualitative analysis: A qualitative assessment will explore each algorithm's design principles and examine their inherent defenses against known quantum attack methods.

3.5. Case Studies and Experiments

This study does not include original case studies or experimental evaluations conducted by the author. Instead, the applicability of post-quantum cryptographic algorithms is measured through a detailed evaluation and analysis of existing case studies, simulations, and experimental data stated in peer-reviewed literature and technical reports.

The evaluation emphasizes on considerate the theoretical efficiency, scalability, and practicality of these algorithms based on documented findings in the field. This approach allows for an informed analysis of the algorithms' potential for real-world deployment, without conducting new experimental work.

3.6. Comparison and Evaluation

3.6.1. Comparison

A comparative analysis will highlight the strengths and weaknesses of each post-quantum cryptographic algorithm, analyzing the competition between security, efficiency, and practicality. This assessment will emphasize on classifying the optimum stability for protection against quantum threats.

3.6.1.1. Key post-quantum cryptographic algorithms

1. Lattice-based cryptography
   - N-th degree truncated polynomial ring (NTRU): Between the first lattice-based cryptographic algorithms, known for its efficiency and resilience.
   - Kyber: A lattice-based encryption scheme chosen as a finalist in the NIST PQC standardization process.
   - Dilithium: A lattice-based digital signature scheme and NIST finalist, valued for its security and efficiency [28], [29].

The (Kyber and Dilithium) Deliver healthy encryption and verification for transmitting complex geospatial data and guarantee protected storage and admission control in GIS platforms.

2. Code-based cryptography
   - McEliece: A venerable code-based cryptosystem with a robust security history, having resisted decades of cryptanalysis.
   - Classic McEliece: A variation of McEliece, chosen as a NIST finalist, which aims to improve efficiency whereas preserving security [30], [31]. It is also ideal for safeguarding large-scale geospatial datasets and avoiding unauthorized decryption of serious GIS information.

3. Multivariate polynomial cryptography
   - Rainbow: A multivariate signature scheme and finalist in the NIST competition, recognized for its rapidity and security properties. The (Rainbow) is also beneficial in validating and collateral spatial inquiries and GIS roadmaps with slight computational overhead.
   - Hidden field equations (HFE): One of the earliest multivariate schemes, laying foundational work for future multivariate cryptographic approaches [32], [33].

4. Hash-based cryptography
   - SPHINCS+: A stateless, hash-based signature scheme and NIST finalist, notable for its resilience to quantum attacks, it also guarantees the integrity and legitimacy of GIS updates and communications, such as map reviews or spatial data sharing.
   - Lamport signatures: An original scheme in hash-based cryptography that serves as a foundational technique for modern hash-based signatures [34], [35].

5. Isogeny-based cryptography
   - Supersingular isogeny key encapsulation (SIKE): An isogeny-based scheme using elliptic curves, selected as a NIST finalist and recognized for its compact key sizes and security [36], It is also right for IoT-enabled GIS systems, where lightweight cryptographic explanations are vital for protected communication between sensors and servers.

3.6.2. Evaluation

1. Lattice-based cryptography
   - NTRU
   - Security: Relies on the hardness of lattice problems, such as the SVP, making it robust against cryptanalysis.
   - Efficiency: Highly efficient and has been a proven encryption algorithm since 1996.
   - Key size: Moderate; public key size is approximately 700 bytes.
   - Signature size: Not applicable, as NTRU is an encryption algorithm.
   - Key operations: Features fast encryption and decryption, supported by efficient polynomial multiplications.
   - Quantum resilience: Demonstrates strong resistance to quantum attacks, with no known efficient quantum algorithms capable of solving the underlying lattice problems.
   - Kyber (NIST finalist)
   - Security: Built on the LWE problem, a cornerstone of lattice-based cryptography.
   - Efficiency: Optimized for both encryption and decryption, offering better performance than traditional schemes.
   - Key size: Public key size is approximately 1,536 bytes.
   - Signature size: Not applicable, as Kyber is an encryption algorithm.
   - Key operations: Outperforms NTRU in key generation and encryption efficiency.
   - Quantum resilience: Highly resistant to quantum attacks due to the inherent difficulty of solving LWE problems with quantum algorithms.
   - Dilithium (NIST Finalist)
   - Security: Secured by the Module Learning with Errors problem, a variant of LWE adapted for modular lattices.
   - Efficiency: A signature scheme offering efficient signing and verification processes.
   - Key size: Public key size is approximately 1,312 bytes.
   - Signature size: Around 2,700 bytes, balancing compactness with security.
   - Key operations: More efficient than many traditional lattice-based signature schemes, with faster performance.

- Quantum resilience: Strongly resilient against quantum attacks, leveraging the well-established hardness of lattice-based problems.

2. Code-Based Cryptography
    - McEliece (Classic McEliece – NIST Finalist)
- Security: Relies on the hardness of decoding general linear codes, a well-established problem in coding theory.
- Efficiency: Offers very efficient encryption and decryption operations.
- Key size: Public key size is significantly large, typically in the range of hundreds of kilobytes to around 1MB.
- Signature size: Not applicable, as McEliece is an encryption algorithm.
- Key operations: While decryption is fast, the large key sizes can make the algorithm less practical for resource-constrained environments or applications requiring frequent key exchanges.
- Quantum resilience: Considered highly secure against quantum attacks, with no substantial progress made by quantum algorithms in breaking its underlying structure.

3. Multivariate Polynomial Cryptography
    - Rainbow (NIST Finalist)
- Security: Based on the difficulty of solving multivariate polynomial equations, an NP-complete problem.
- Efficiency: Highly efficient in terms of signing and verification speed, though it requires larger key sizes.
- Key Size: Public key size ranges between 50 and 100 KB, depending on the parameter set.
- Signature size: Approximately 66 bytes, making it compact for a signature scheme.
- Key operations: Features fast signing and verification processes, suitable for applications needing rapid authentication.
- Quantum resilience: Multivariate schemes, including Rainbow, are generally regarded as secure against quantum attacks, but ongoing research continues to evaluate their robustness.
    - HFE
- Security: Relies on the complexity of HFE, a problem considered hard to solve even with quantum computing advancements.
- Efficiency: Offers greater efficiency than a rainbow in some configurations but also requires large key sizes.

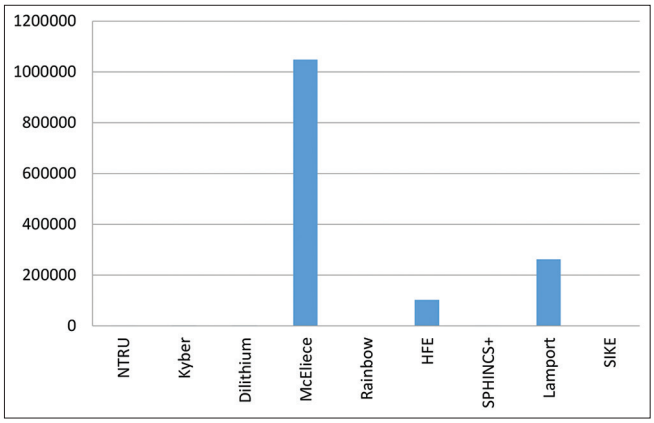


Chart 1. Key sizes in byes versus algorithms

- Key size: Typically, around 100 KB or more.
- Signature size: Variable depending on parameters but can be relatively small compared to other multivariate schemes.
- Key operations: Provides fast signing and verification with some computational overhead during key generation.
- Quantum resilience: Considered secure against currently known quantum attacks, though, such as a rainbow, its resilience is subject to ongoing evaluation.

4. Hash-based cryptography
    - SPHINCS+ (NIST Finalist)
- Security: Relies on well-established hash functions, making it inherently resistant to quantum attacks.
- Efficiency: Implements a stateless hash-based signature approach; while secure, it is less efficient compared to other post-quantum schemes.
- Key size: Compact private keys (~32 bytes).
- Signature size: Relatively large, approximately 41 KB, due to the use of Merkle tree structures.
- Key operations: Signing is slower because of tree traversal, but verification is notably faster than many alternative schemes.
- Quantum resilience: Exceptionally strong; hash-based cryptographic methods are naturally resistant to Grover's algorithm and other quantum-based attacks.
    - Lamport signatures
- Security: A simple hash-based cryptographic system offering basic security rooted in the hardness of hash functions.
- Efficiency: Highly inefficient, with extremely large key sizes and a 1-time-use requirement for keys.

Table 1: Summary of cryptographic algorithms comparison

Algorithm	Security basis	Key size	Signature size	Efficiency	Quantum resilience
N-th degree truncated polynomial ring	Lattice (shortest vector problem)	~700 bytes	N/A	Fast	Strong
Kyber	Lattice (learning with errors)	~1,536 bytes	N/A	Very fast	Strong
Dilithium	Lattice (module learning with errors)	~1,312 bytes	~2,700 bytes	Fast	Strong
McEliece	Code-based (linear codes)	~1 MB	N/A	Fast encryption, slow key generation	Very strong
Rainbow	Multivariate polynomial equations	~50–100 KB	~66 bytes	Fast	Good, but more research needed
Hidden field equations	Hidden field equations	~100 KB	Medium	Fast	Strong
SPHINCS+	Hash-based	~32 bytes	~41 KB	Secure, but slow signing	Very strong
Lamport	Hash-based	Very large	Large	Impractical	Very strong
Supersingular isogeny key encapsulation	Isogenies on elliptic curves	~330 bytes	N/A	Slow	Promising, but newer

Table 2: Algorithm security-based comparison

Algorithm	Security basis	Key features	Resilience to quantum attacks
N-th degree truncated polynomial ring	Lattice (SVP)	Efficient for encryption and key exchange.	Relies on the SVP. Hard for both classical and quantum computers.
Kyber	Lattice (LWE)	Compact keys, highly efficient. NIST finalist.	Based on the LWE problem. Quantum-safe.
Dilithium	Lattice (M-LWE)	Designed for digital signatures. NIST finalist.	Variant of LWE optimized for signatures. Quantum-safe.
McEliece	Code-based (Linear codes)	Large public keys but very fast encryption.	Based on hard decoding of random linear codes. Resistant to quantum attacks.
Rainbow	Multivariate polynomial equations	Multivariate public-key signatures.	Based on solving polynomial systems, but currently vulnerable to structural attacks.
Hidden field equations	Hidden field equations	Compact signatures but slower verification.	Relies on hidden field structure equations. Generally, quantum-safe.
SPHINCS+	Hash-based	Stateless digital signatures.	Hash-based security. Resistant to quantum pre-image attacks.
Lamport	Hash-based	Early hash-based signature scheme, simple but practical.	Secure if hash functions are strong. Limited usability for single-use scenarios.
Supersingular isogeny key encapsulation	Isogenies on elliptic curves	Extremely small key sizes. NIST candidate but recently broken.	Relies on isogeny problems, but a recent quantum attack compromised its viability.

SVP: Shortest vector problem, LWE: Learning with errors, M-LWE: Module learning with errors

Table 3: Algorithm signature size comparison

Algorithm	Signature Size	Signature size details
N-th degree truncated polynomial ring	N/A	Small (Around~600 bytes for experiments; not primary focus)
Kyber	N/A	Not designed for signatures (Only key exchange/encryption)
Dilithium	~2,700 bytes	~2,700 bytes
McEliece	N/A	Large (Experimental schemes suggest~135 KB for signatures)
Rainbow	~66 bytes	~66 bytes
Hidden field equations	Medium	Medium (Exact size varies with parameters; generally, in KB range)
SPHINCS+	~41 KB	~41 KB
Lamport	Large	Large (~262 KB or more depending on security level)
Supersingular isogeny key encapsulation	N/A	Experimental only (Compact signatures around~300 bytes in research setups)

- Key size: Extremely large, often hundreds of megabytes, making it impractical for most applications.
- Signature size: Substantial, though smaller than SPHINCS+ signatures.
- Key operations: Each signature requires a new key, leading to significant overhead and limiting practical usability.
- Quantum resilience: Fully quantum-resistant, but its size and single-use nature render it unsuitable for widespread adoption.

5. Isogeny-based cryptography
    - SIKE – NIST Finalist

- Security: Relies on the difficulty of computing isogenies between supersingular elliptic curves, a relatively new cryptographic hardness assumption.
- Efficiency: Less efficient than lattice-based and code-based schemes, particularly in terms of computation speed.
- Key size: Compact public keys (~330 bytes), making it attractive for environments where storage or transmission bandwidth is limited.
- Signature size: Not applicable, as SIKE is a key encapsulation mechanism rather than a signature scheme.
- Key operations: Key generation and encapsulation processes are slower compared to other post-quantum cryptographic algorithms.
- Quantum resilience: While considered quantum-resistant, its underlying security assumptions are newer and remain under active scrutiny, unlike more established systems, such as lattice-based cryptography.

Table 4: Algorithm efficiency comparison

Algorithm	Efficiency
N-th degree truncated polynomial ring	Fast
Kyber	Very fast
Dilithium	Fast
McEliece	Fast encryption, slow key generation
Rainbow	Fast
Hidden field equations	Fast
SPHINCS+	Secure, but slow signing
Lamport	Impractical
Supersingular isogeny key encapsulation	Slow

The examination of the algorithms is offered through various tables and a chart in this study, each presents an exceptional viewpoint as follows: Table 1 provides a wide range of comparison for each type of algorithm, while Table 2 focuses on their security features. Table 3 studies the signature sizes of the algorithms, and Table 4 appraises their general efficiency. Table 5 highlights the flexibility of the algorithms to quantum attacks, and Table 6 summarizes the basis of security for each. Additionally, Chart 1: presents a visual comparison of the key sizes for each algorithm in bytes, presenting valuable understanding into their scalability and practicality for various use cases.

Table 5: Algorithm quantum resilience comparison

Algorithm	Quantum resilience	Details
N-th degree truncated polynomial ring	Strong	NTRU is considered strong against quantum attacks due to its reliance on lattice-based problems, which are not easily solvable using quantum algorithms, such as Shor's algorithm.
Kyber	Strong	Kyber, another lattice-based algorithm, is highly resilient against quantum computing attacks, particularly those targeting number-theoretic problems.
Dilithium	Strong	Dilithium, which uses lattice-based cryptography (M-LWE), also offers strong resilience against quantum threats, similar to Kyber and NTRU.
McEliece	Very strong	McEliece is very strong against quantum attacks due to its foundation in coding theory (specifically, the hardness of decoding random linear codes), which is not susceptible to quantum algorithms.
Rainbow	Good, but more research needed	Rainbow, a multivariate polynomial-based signature scheme, shows good resilience but is still being researched to determine its robustness against quantum attacks.
HFE	Strong	HFE are generally strong against quantum attacks, but their implementation complexity and performance can be limiting.
SPHINCS+	Very strong	SPHINCS+is a very strong hash-based signature scheme, resistant to quantum attacks, and its security is based on the collision resistance of hash functions, which is not impacted by quantum computers.
Lamport	Very strong	Lamport signatures, being based on hash functions, are very strong in the context of quantum resilience, as they rely on the collision resistance of cryptographic hashes.
SIKE	Promising, but newer	SIKE is promising, but it's a newer approach to post-quantum cryptography and requires more research to establish its long-term resilience against quantum threats.

M-LWE: Module learning with errors, HFE: Hidden field equations, SIKE: Supersingular isogeny key exchange

Table 6: Algorithm security basis comparison

Algorithm	Security basis	Explanation
N-th degree truncated polynomial ring	Lattice (SVP)	NTRU is based on the SVP in lattice theory. SVP is considered difficult to solve even for quantum computers, providing strong security against quantum attacks.
Kyber	Lattice (LWE)	Kyber is based on the LWE problem, a well-studied problem in lattice-based cryptography. It is efficient and resistant to quantum attacks.
Dilithium	Lattice (M-LWE)	Dilithium uses M-LWE, a variation of LWE with improved efficiency, particularly in digital signatures. Like LWE, M-LWE is resistant to quantum attacks.
McEliece	Code-based (Linear codes)	McEliece is based on error-correcting codes, specifically decoding random linear codes, which is considered very difficult even for quantum computers.
Rainbow	Multivariate polynomial equations	Rainbow uses multivariate polynomial equations to construct its signature scheme. It relies on the difficulty of solving systems of multivariate polynomials, which is believed to be hard for both classical and quantum computers.
HFE	Hidden field equations	HFE is a public-key cryptosystem based on the difficulty of solving systems of equations in finite fields. It is considered secure, but less studied compared to other quantum-resistant schemes.
SPHINCS+	Hash-based	SPHINCS+is a hash-based signature scheme, utilizing the security of hash functions. Since hash functions are believed to be quantum-secure (with only a quadratic speedup from quantum algorithms), SPHINCS+offers strong security against quantum computers.
Lamport	Hash-based	Lamport is a hash-based signature scheme that uses the collision resistance of hash functions for security. Like SPHINCS+, it is considered quantum-secure.
SIKE	Isogenies on elliptic curves	SIKE relies on the difficulty of finding isogenies (special mappings) between supersingular elliptic curves. While promising, it is still under research, and quantum resilience is still being evaluated.

M-LWE: Module learning with errors, HFE: Hidden field equations, SIKE: Supersingular isogeny key exchange, SVP: Shortest vector problem, LWE: Learning with errors

4. DISCUSSION

4.1. Lattice-Based Algorithms (Kyber, Dilithium)

These algorithms are highly effective and feature fairly small key sizes, making them compatible for a comprehensive sequence of applications, including constrained situations. They are between the most secure choices existing, backed by extensive research and proven mathematical basics in the post-quantum cryptographic arena.

4.2. Code-Based Algorithms (McEliece)

While McEliece is famous for its strong security, the impossibly large key sizes position challenges for widespread adoption, primarily in systems with inadequate storage or transmission abilities. However, its long-standing confrontation to cryptography, which makes it a reliable choice in extremely sensitive applications.

4.3. Hash-Based Algorithms (SPHINCS +)

These algorithms offer supreme security, particularly for applications requiring a long-term battle against quantum attacks. However, the inadequacy of signing processes and larger signature sizes may limit their usability in performance-critical systems.

4.4. Isogeny-Based Algorithms (SIKE)

SIKE stands out for its compacted key sizes, which are beneficial in situations with bandwidth or storage limitations. However, its slower performance and the quite nascent nature of its security expectations involve further scrutiny and research before widespread adoption.

5. CONCLUSION

The rapid advancements in quantum computing present considerable threats to traditional cryptographic systems, highlighting the vital need for a modification to PQC. This paper has analyzed numerous promising post-quantum cryptographic algorithms, including lattice-based, code-based, multivariate polynomial, hash-based, and isogeny-based approaches. Key issues measured in the evaluation include security, efficiency, quantum resilience, key size, and computational performance. Quantum computing is also playing a key role in GIS systems hinging on protected and efficient cryptographic methods to manage and allocate spatial data. Implementing post-quantum algorithms guarantees these systems remain strong as quantum computing abilities rise.

The conclusions designate that lattice-based and hash-based cryptography reveal brilliant confrontation to quantum

attacks, making them robust applicants for safeguarding upcoming infrastructures. Algorithms such as Kyber and SPHINCS+ excel in matching quantum resistance with practical deliberations, such as key size and computational efficiency. However, contests remain, particularly in enhancing key generation periods and certifying competence in real-world applications. In conclusion, the insights from this study contribute to the developing field of PQC, arranging groundwork for the development of secure communication systems in quantum time. As quantum computing remains in development, cryptographic systems must adjust to meet these emerging contexts. The algorithms discussed in this paper provide a solid initiation point for designing robust, quantum-resistant cryptographic solutions to safeguard upcoming communication systems.

The future direction of this paper should focus on the following, which we couldn't cover them in this study, which are the:

- Optimizing post-quantum algorithms: Research could focus on improving the efficiency and scalability of PQC algorithms.
- Broader geographic analysis: Investigate how specific geopolitical factors influence the adoption of quantum-resistant algorithms.
- Standardization efforts: discover association opportunities with international bodies.
- Interdisciplinary approaches: Influence fields such as artificial intelligence or machine learning to enhance cryptographic resilience in quantum-ready environments.

REFERENCES

[1] P. Williams, I. K. Dutta, H. Daoud and M. Bayoumi. "A survey on security in internet of things with a focus on the impact of emerging technologies". *Internet of Things*, vol. 19, p. 100564, 2022.

[2] R. Rietsche, C. Dremel, S. Bosch, L. Steinacker, M. Meckel and J. M. Leimeister. "Quantum computing". *Electronic Markets*, vol. 32, no. 4, pp. 2525-2536, 2022.

[3] J. J. Tom, N. P. Anebo, B. A. Onyekwelu, A. Wilfred and R. E. Eyo. "Quantum computers and algorithms: A threat to classical cryptographic systems". *International Journal of Engineering and Advanced Technology*, vol. 12, no. 5, pp. 25-38, 2023.

[4] A. Naik, E. Yeniaras, G. Hellstern, G. Prasad and S. K. L. P. Vishwakarma. "From portfolio optimization to quantum blockchain and security: A systematic review of quantum computing in finance". *arXiv [cs.CR]*, 2023.

[5] N. Sood. "Cryptography in post quantum computing era". *SSRN Electronic Journal*, 2024. Available: Available at SSRN 4705470 [Last accessed on 2025 Jan 23].

[6] M. Kumar and P. Pattnaik. "Post Quantum Cryptography (PQC)-an Overview". In: *2020 IEEE High Performance Extreme Computing Conference* (*HPEC*). IEEE, 2020.

[7] H. Andås. "Emerging technology trends for defence and security". 2020.

[8] M. Kirshner. "*Achieving Holistic Interoperability with Model-Based Systems Engineering*". Ph.D. dissertation, The University of Arizona, 2023.

[9] J. Hekkala, M. Muurman, K. Halunen and V. Vallivaara. "Implementing post-quantum cryptography for developers". *SN Computer Science*, vol. 4, no. 4, p. 365, 2023.

[10] Y. Baseri, V. Chouhan and A. Ghorbani. "Cybersecurity in the quantum era: Assessing the impact of quantum computing on infrastructure". *arXiv [cs.CR]*, 2024.

[11] I. Gunawan, S. Sumarno, H. S. Tambunan, E. Irawan, H. Qurniawan and D. Hartama. "Combination of caesar cipher algorithm and rivest shamir adleman algorithm for securing document files and text messages". *Journal of Physics: Conference Series*, vol. 1255, p. 012077, 2019.

[12] Z. Hu, B. Liu, X. Ren and Y. Tang. "Analysis and Implementation of the Enigma Machine". In *2022 International Conference on Big Data, Information and Computer Network* (*BDICN*). IEEE, 2022.

[13] R. Slayton. "*Democratizing Cryptography: The Work of Whitfield Diffie and Martin Hellman*". ACM, United States, 2022.

[14] Y. Zhang, A. Liu, C. Liu, B. Ai and X. Zhang. "A track initiation algorithm using residual threshold for shore-based radar in heavy clutter environments". *Journal of Marine Science and Engineering*, vol. 8, no. 8, p. 614, 2020.

[15] S. Solanki, S. Sharma and A. Yahya. "Quantum algorithms: Unleashing the power of quantum computing". *OORJA - International Journal of Management & IT*, vol. 21, no. 1, p. 28, 2023.

[16] M. R. Habibi, S. Golestan, A. Soltanmanesh, J. M. Guerrero and J. C. Vasquez. "Power and energy applications based on quantum computing: The possible potentials of grover's algorithm". *Electronics*, vol. 11, no. 18, p. 2919, 2022.

[17] K. F. Hasan, L. Simpson, M. Baee, C. Islam, Z. Rahman, W. Armstrong, P. Gauravaram and M. McKague. "A framework for migrating to post-quantum cryptography: Security dependency analysis and case studies". *IEEE Access*, vol. 12, pp. 23427-23450, 2024.

[18] M. A. Khan, S. Javaid, S. A. H. Mohsan, M. Tanveer and I. Ullah. "Future-proofing security for UAVs with post-quantum cryptography: A review". *IEEE Open Journal of the Communications Society*, vol. 5, pp. 6849-6871, 2024.

[19] P. K. Pradhan, S. Rakshit and S. Datta. "Lattice Based Cryptography: Its Applications, Areas of Interest & Future Scope". In *2019 3rd International Conference on Computing Methodologies and Communication* (*ICCMC*). IEEE, 2019.

[20] G. Nookala. "Post-quantum cryptography: Preparing for a new era of data encryption". *MZ Computing Journal*, vol. 5, no. 2, p. 012077, 2024.

[21] A. Kichna and A. Farchane. "Secure and efficient code-based cryptography for multi-party computation and digital signatures". *Computer Sciences and Mathematics Forum*, vol. 1, p. 1, 2023.

[22] R. Kuang and M. Perepechaenko. "Optimization of the multivariate polynomial public key for quantum safe digital signature". *Scientific Reports*, vol. 13, no. 1, p. 6363, 2023.

[23] M. Singh, S. K. Singh, S. Kumar, M. Preet, V. Arya and B. B. Gupta. "Quantum-resilient cryptographic primitives: An innovative modular hash learning algorithm to enhanced security in the quantum era". *Research Square,* 2024.

[24] A. Karakaya and A. Ulu. “A survey on post-quantum based approaches for edge computing security”. *Wiley Interdisciplinary Reviews: Computational Statistics*, vol. 16, no. 1, p. e1644, 2024.

[25] K. S. Roy and H. K. Kalita. “A survey on post-quantum cryptography for constrained devices”. *International Journal of Applied Engineering Research*, vol. 14, no. 11, pp. 2608-2615, 2019.

[26] M. Das, A. Nag, M. Hassan, A. Santra, N. Chand, F. Yasmin, A. Sinha, A. K. Bairagi and A. Alkhayyat. “Synergy of 6G technology and IoT networks for transformative applications”. *International Journal of Communication Systems*, vol. 37, no. 14, p. e5869, 2024.

[27] J. N. Pelton and S. Madry. “*Space Systems, Quantum Computers, Big Data and Sustainability: New Tools for the United Nations Sustainable Development Goals*”. CRC Press, United States, pp. 53-104, 2024.

[28] J. Hoffstein. “NTRU: A ring based public key cryptosystem”. In: *Algorithmic Number Theory* (*ANTS III*). Springer, Berlin, Heidelberg, 1998.

[29] V. Lyubashevsky, C. Peikert and O. Regev. “On ideal lattices and learning with errors over rings”. In: *Advances in Cryptology-EUROCRYPT 2010*. Springer, Berlin, Heidelberg, pp. 29-48, 2010.

[30] R. J. McEliece. “A public-key cryptosystem based on algebraic”. *Coding Thv*, vol. 4244, pp. 114-116, 1978.

[31] D. J. Bernstein, T. Lange and C. Peters. “Attacking and defending the McEliece cryptosystem”. In: *Post-Quantum Cryptography*. Springer, Berlin, Heidelberg, pp. 31-46, 2008.

[32] J. Patarin. “Hidden Fields Equations (HFE) and Isomorphisms of Polynomials (IP): Two New Families of Asymmetric Algorithms”. In: *International Conference on the Theory and Applications of Cryptographic Techniques*. Springer, Berlin, Heidelberg, 1996.

[33] J. Ding and D. Schmidt. “Rainbow, a new multivariable polynomial signature scheme”. In: *International Conference on Applied Cryptography and Network Security*, Springer, Berlin, Heidelberg, 2005.

[34] L. Lamport. “Constructing digital signatures from a one way function”. 1979. Available from: https://www.microsoft.com/en-us/research/publication/constructing-digital-signatures-one-way-function/ [Last accessed on 2025 Jan 22].

[35] D. J. Bernstein, D. Hopwood, A. Hülsing, T. Lange, R. Niederhagen, L. Papachristodoulou, M. Schneider, P. Schwabe and Z. Wilcox-O’Hearn. “SPHINCS: Practical Stateless Hash-based Signatures”. In: *Annual International Conference on the Theory and Applications of Cryptographic Techniques*, Springer, Berlin, Heidelberg, 2015.

[36] D. Jao and L. De Feo. “Towards Quantum-resistant Cryptosystems from Supersingular Elliptic Curve Isogenies”. In *Post-Quantum Cryptography: 4th International Workshop*. Springer, Berlin, Heidelberg, 2011.

ORIGINAL RESEARCH ARTICLE UHD JOURNAL OF SCIENCE AND TECHNOLOGY

The Clinical Neurological Manifestations of Patients Diagnosed with Carpal Tunnel Syndrome

Omar Hussein Shareef[1], Shorsh Ahmed Mohhammed[2], Hemn Mohammed Gharib[2]

[1]Department of Community Health Nursing, College of Nursing, University of Sulaimani, [2]Sulaimani General Directorate of Health, Sulaymaniyah, Iraq.

ABSTRACT

Background: Carpal tunnel syndrome (CTS) is a condition, in which the median nerve becomes pressed or squeezed at the wrist. This causes pain and numbness in the fingers. Therefore, a neurological study is crucial to assess the condition. **Objectives:** The objective of this study was to assess the neurological manifestations of CTS and their association with demographic and clinical features from October 2022 to March 2023. **Materials and Methods:** A quantitative study was carried out over the period of 5 months by prospectively selecting and enrolling 100 CTS patients with a confirmed diagnosis. The CTS assessment questionnaire was modified and patients consented to the study before the data collection. **Results:** Adults aged 35–44 were the dominant group and the disease was found in females 10 times more than males. The least assigned symptoms were tingling and numbness in the little finger (4%) and neck pain 22%. All the patients with CTS presented with severe levels of CTS. Statistically significant associations were found between occupations, duration of the disease, affected side, other chronic diseases, and the prevalence of the symptoms at $P \leq 0.05$. Self-management to sub-side pain and numbness had crucial impact on reducing the symptoms ($P \leq 0.05$). **Conclusion:** The prevalence of the neurological symptoms varied depending on the sociodemographic and clinical features. Self-management had a significant positive impact on reducing some of the neurological symptoms, such as pain in the wrist at night and tingling and numbness in the morning.

Index Terms: Carpus Tunnel Syndrome, Carpal Tunnel Syndrome, Neurological Symptoms

1. INTRODUCTION

A group of symptoms caused by compression of the median nerve at the carpal tunnel is known as carpal tunnel syndrome (CTS). Pain, numbness, or tingling on the anterior surface of the index, middle, or radial half of the ring finger are signs of compression of the median nerve. It is frequently linked to hand grip weakness or nighttime symptoms, including numbness and pain in the hands or arms [1].

Access this article online

DOI: 10.21928/uhdjst.v9n1y2025.pp29-33

E-ISSN: 2521-4217

P-ISSN: 2521-4209

The pathophysiological mechanism linked to CTS is the median nerve entrapment in the wrist, which causes compression of the nerve, resulting in paresthesia, numbness, and hand muscle weakness that can be diagnosed by the nerve conduction study [2].

The disease is very common in the general population; at a point, the morbidity is rising with the advancement of modern life. Since CTS is a common illness, research on its etiology is crucial to improving the quality of life among patients. Individual characteristics such as age, sex, diabetes, hypothyroidism, obesity, tobacco, injury, and occupational factors are among the risk factors. However, the majority of the studies were conducted on western societies [3]. A multifactorial etiology has been suggested by the association between CTS and occupational i.e., repetitive and forceful work such as gripping and vibrations [4].

Corresponding author's e-mail: Dr. Omar Hussein Shareef, Department of Community Health Nursing, College of Nursing, University of Sulaimani, Sulaymaniyah, Iraq. E-mail: omar.shareef@univsul.edu.iq.

Received: 19-07-2024 Accepted: 09-09-2024 Published: 20-02-2025

The links between CTS and work-related psychological aspects at work are still unclear, despite an increasing number of studies on occupational stress [5,6]. Because of increased exposure to repetitive movements and a lack of recovery time, lean production systems and the corresponding new work organization systems are thought to intensify work pace and job demand and, as a result, increase the risk of musculoskeletal disorders. Some epidemiological studies indicate associations between a higher risk of CTS and the characteristics of the workplace [7]. Despite the available evidence of high certainty about work□related physical risk factors for CTS, there still might be a lack of awareness in clinical care for prevention [4].

The current study aims to assess the neurological manifestations of CTS and their association with demographic and clinical features.

2. MATERIALS AND METHODS

2.1. Study Design and Population

A quantitative study of cross-sectional design was carried out over the period of 5 months starting from October 2022 to March 2023. The process of the study was declared to the participants and verbal consent was taken from the patients following the ethical guarantee from the Ethical Committee for all Medical Colleges (College of Medicine, College of Dentistry, College of Pharmacy, and College of Nursing) at the University of Sulaimani.

2.2. Sampling Procedures

An accidental sampling procedure was applied to interview 100 patients with CTS who consented to the study and had the characteristics of inclusion criteria.

2.3. Inclusion and Exclusion Criteria

Patients diagnosed with CTS were included after their consent. Patients with non-confirmed diagnoses and patients who did not wish to participate were excluded.

2.4. Diagnostic Tools

The diagnostic tool: The diagnosis of this disorder has been performed by the patients' chief complaint, and provocative physical examination procedures such as Phalen's sign and Tinel's sign. Phalen's test is a series of hand motions and positions that will make the hands or fingers feel numb or tingly if a client has CTS. Tinel's sign is a tingling feeling when the clinician taps the client's skin over an affected nerve [5].

2.5. Data Collection

Data were gathered by conducting interviews with using a modified questionnaire in Sulaimani Rehabilitation and Physiotherapy Center Qanat Street.

2.6. Data Analysis

The Statistical Package for Social Science version 25.0 was used to analyze the data. Several statistical tests, including, percentages, and frequencies, are among the descriptive statistics. In the inferential statistics, Chi-square and Fisher's exact tests were used to analyze the data. The alpha level value of 0.05 was used as the cutoff for statistical significance and 0.001 for high statistical significance.

3. RESULTS

Table 1 shows the sociodemographic characteristics of the study participants, the age group between 35 and 44 years old comprises 54% of the total achieved sample. According to the findings of the present study, females were affected by CTS when compared to males (91% female and 9% male). In addition, the majority of the patients were married (69%), a quarter of the total sample was illiterate, and eight patients completed secondary school. In terms of occupation, which may have a significant impact on the prevalence of the disease, the majority of the participants were housewives, and the proportion of the patients who retired from jobs was only 1%, besides patients who were employed comprised 34% of the total sample.

Table 2 describes the clinical characteristics of the participants, the proportion of the patients who never consumed alcohol was 99%, however, 18% of them were smokers. The majority of the patients had the disease for more than 1 year, which is 71%, and the rest have it for <1 year. Regarding the side that was affected by the neurological symptoms, right arm symptoms were doubled when compared to the left side, while patients who had symptoms on both arms had. Patients who have not been diagnosed with other chronic diseases were nearly twice as many as those with other chronic diseases (65% vs. 35%, respectively). Type 2 diabetes mellitus was the most prevalent chronic disease besides CTS (11%), while only 2% of them had hypertension and gout disease. Finally, ten patients (10%) of the total sample were living with more than five comorbidities.

Table 3 shows the neurological clinical symptoms among the patients. Seventy-seven percent of the patients had pain in

TABLE 1: Sociodemographic characteristics of the study participants

Sociodemographic features	Frequency	Percent
Age groups		
35–44	52	52.0
45–54	29	29.0
55–64	12	12.0
65–74	6	6.0
75–84	1	1.0
Gender		
Male	9	9.0
Female	91	91.0
Marital Status		
Single	15	15.0
Married	69	69.0
Divorced	4	4.0
Widow or Widower	12	12.0
Level of education		
Illiterate	25	25.0
Primary school	20	20.0
Secondary school	8	8.0
Diploma	26	26.0
University and higher	21	21.0
Occupation		
Employed	34	34.0
Self-employed	5	5.0
Housewife	47	47.0
Retired	1	1.0
Jobless	13	13.0

TABLE 2: Clinical characteristics among the participants

Clinical data	n.	(%)
Alcohol consumption		
Yes	1	1.0
No	99	99.0
Total	100	100.0
Smoking status		
Smoker	18	18.0
Non-Smoker	77	77.0
Ex-Smoker	5	5.0
Duration of CTS		
<1 year	29	29.0
More than 1 year	71	71.0
Side of CTS		
Right side	55	55.0
Left Side	29	29.0
Both sides	16	16.0
Other chronic diseases		
Yes	35	35.0
No	65	65.0
Types of other chronic diseases		
Gout	2	2.0
Hypothyroid	4	4.0
DM	11	11.0
Rheumatoid arthritis	6	6.0
HTN	2	2.0
more than 5 diseases	10	10.0
Total	35	35.0
No Other diseases	65	65.0
Total	100	100.0

TABLE 3: Prevalence of neurological symptoms among the patients

Neurological symptoms	n.	(%)
Pain in wrist at night	77	77.0
Tingling and numbness at night	78	78.0
Tingling and numbness in morning	72	72.0
Tingling and numbness in little finger	4	4.0
Tingling and numbness when grasping something	54	54.0
Neck pain	22	22.0

their wrists at night, besides tingling and numbness, which accounted for 78% of the total sample. Morning symptoms also comprised 72% of the participants, particularly tingling and numbness. In addition, more than half of the patients 54% reported tingling and numbness when grasping something.

Regarding specific areas of the symptoms, only 4 percent complained of tingling and numbness in the little finger, and neck pain was presented in 22%.

In Table 4, a statistically significant relationship was found between the occurrence of tingling and numbness in the morning and the patients' occupation at $P = 0.05$. To exemplify, this symptom was reported among 100% of patients who were self-employed and more than 80% of housewife patients who work repetitively handy work at home. Moreover, all the patients who were retired from their jobs complained of the same symptoms.

Table 5 illustrates a statistically highly significant association between the clinical features of the participants and the CTS symptoms. For example, 82% of the patients who had the disease for more than 1 year presented with tingling and numbness in the morning when compared to those who had a history of <1 year which only 48% complained with this symptom at $P = 0.001$.

Tingling and numbness at night were found among those who had other chronic diseases by almost 91%. Patients with rheumatoid arthritis (RA), HTN, and comorbidity (more than five diseases) presented without tingling and numbness in the morning $P = 0.016$.

Table 6 discloses the level of CTS symptoms among the patients, by which it has being found that 100% of the study sample presented with severe levels of CTS symptoms.

4. DISCUSSION

In the current study, most of the CTS were female and they were doing housekeeping in terms of occupation. This is totally in line with the finding of a study by Hong *et al*., 2022 in China, by which female comprises the majority of the cases when compared to males (72 females and 2 males). Among the 72 females, 29 were housekeepers and 18 were farmers. In another study that aimed to explore the characteristics of CTS patients, it has being found that 63% of the cases are female and 37% – male [8].

TABLE 4: Association between patients' occupation and the CTS symptom

Occupation	Tingling and numbness in morning		Total
	Yes	No	
	n (%)	n (%)	
Employed	21 (61.7)	13 (38.3)	34
Self-employed	5 (100)	0 (0.0)	5
Housewife	38 (80.9)	9 (19.1)	47
Retired	1 (100)	0 (0.0)	1
Jobless	7 (53.8)	6 (46.2)	13
Total	72 (72)	28 (28)	100
FET*=7.5		*P*=0.05	

TABLE 5: Association clinical features of the patients and the CTS symptom

Duration of CTS	Tingling and numbness in morning		Total
	Yes n (%)	No n (%)	
<1 year	14 (48)	15 (52)	29
More than 1 year	58 (82)	13 (12)	71
Total	72 (72)	28 (28)	100
Pearson Chi-Square=11.4		*P*=0.001	
Side of CTS	**Tingling and numbness in little finger**		**Total**
	Yes	**No**	
Right side	0 (0.0)	55 (100)	55
Left side	1 (3.5)	28 (96.5)	29
Both sides	3 (18.7)	13 (81.3)	16
Total	4 (4)	96 (96)	100
FET=11.3		*P*=0.03	
Other chronic diseases	**Tingling and numbness at night**		**Total**
	Yes	**No**	
Yes	31 (91)	3 (9)	34
No	47 (71)	19 (29)	66
Total	78 (78)	22 (22)	100
Pearson Chi-square=5.21		*P*=0.02	
Types of other Chronic Diseases	**Tingling and numbness in morning**		**Total**
	Yes	**No**	
Gout	1 (50)	1 (50)	2
Hypothyroid	3 (75)	1 (25)	4
DM	8 (73)	3 (27)	11
Rheumatoid arthritis	6 (100)	0 (0.0)	6
HTN	2 (100)	0 (0.0)	2
more than 5 diseases	10 (100)	0 (0.0)	10
Total	30 (30)	5 (5)	35
FET=6.9		*P*=0.016	

In our study, the proportion of the right side wrist affected with CTS was 55%, the left side – 29%, and both wrists – 16%, but these results are noticed in other studies differently. For instance, Hong *et al*. found only 30% were affected with the right side, 15% – left wrist, and 55% having CTS in both wrists [8].

In the current study, a significant relationship was observed between the occupations and the occurrences of CTS. Similarly, in several studies, the role of occupations was examined among workers and concluded with the significant effects of occupation on the prevalence of CTS [9]–[11]

In this study, moreover, there is a great impact of patients' clinical data on the intensity and variation of the neurological symptoms, For example, the duration of the disease, the affected side of CTS, and the patients' other chronic diseases had a critical impact on the symptoms, patients with type 2 diabetes mellitus presented with severe neurological symptoms, especially tingling and numbness. The results of a case–control study were in line with our study presenting an association between diabetes mellitus and CTS symptoms [12]. In addition, a study of the assessment of the incidence and severity of CTS among RA patients found the same curtailment impact of RA on the incidence and symptoms of CTS [7], [12].

In our study, the level of CTS symptom intensity was analyzed by the perceived CTS questionnaire. Following the scoring, the study sample presented with severe CTS symptoms and no one was found with mild symptoms. To discuss this finding, no studies were found clearly with the same results. However, there is a substantial amount of literature highlighting the causes of the severity of CTS symptoms including, nerve injury, duration of the disease, side of the arm, occupation, and other comorbidities [13]–[16].

TABLE 6: The level of CTS symptoms among the study sample

Level of CTS intensity	Frequency	Percent
Mild CTS symptoms	0	0.0
Sever CTS symptoms	100	100.0

5. CONCLUSION

The study discovered that patients with CTS mainly experience neurological symptoms that interfere with their daily activities. Tingling and numbness at night were the most common symptoms among the patients. Several factors can aggregate the symptoms, including the duration of the disease and other comorbidities and their occupation.

6. ACKNOWLEDGMENT

It is acknowledged to all patients involved in this study.

7. ETHICAL CONSIDERATIONS

This study was approved by the Scientific Committee of the Basic Sciences Department at College of Nursing. Then, Ethical approval was guaranteed from the Ethical Committee for the Medical Colleges at University of Sulaimani. In addition, patients were informed about the objectives of the study and their independent participation. They also provided consent to the current study.

8. CONFLICTS OF INTERESTS

The author declared that they have no conflicts of interest.

9. FUNDING

No financial support was provided for the study.

REFERENCES

[1] C. F. De-Las Peñas, R. Ortega-Santiago, A. I. de la Llave-Rincón, A. Martínez-Perez, H. F. S. Díaz, J. Martínez-Martín, J. A. Pareja and M. L. Cuadrado-Pérez. "Manual physical therapy versus surgery for carpal tunnel syndrome: A randomized parallel-group trial". *The Journal of Pain*, vol. 16, no. 11, pp. 1087-1094, 2015.

[2] P. K. Srikanteswara, J. D. Cheluvaiah, J. B. Agadi and K. Nagaraj. "The relationship between nerve conduction study and clinical grading of carpal tunnel syndrome". *Journal of Clinical and Diagnostic Research*, vol. 10, pp. 13-18, 2016.

[3] M. Sonohata, T. Tsuruta, H. Mine, A. Asami, H. Ishii, K. Tsunoda and M. Mawatari. "The effect of carpal tunnel release on neuropathic pain in carpal tunnel syndrome". *Pain Research and Management*, vol. 2017, p. 8098473, 2017.

[4] M. Riccò, S. Cattani and C. Signorelli. "Personal risk factors for carpal tunnel syndrome in female visual display unit workers". *International Journal of Occupational Medicine and Environmental Health*, vol. 29, no. 6, pp. 927-936, 2016.

[5] G. Ntani, K. T. Palmer, C. Linaker, E. Clare Harris, R. Van der Star, C. Cooper and D. Coggon. "Symptoms, signs and nerve conduction velocities in patients with suspected carpal tunnel syndrome". *BMC Musculoskeletal Disorders*, vol. 14, p. 242, 2013.

[6] T. Koukoulaki. "The impact of lean production on musculoskeletal and psychosocial risks: An examination of sociotechnical trends over 20 years". *Applied Ergonomics*, vol. 45, no. 2, pp. 198-212, 2014.

[7] A. Dabbagh, J. C. Macdermid, J. Yong, L. G. Macedo and T. L. Packham. "Diagnosing carpal tunnel syndrome: Diagnostic test accuracy of scales, questionnaires, and hand symptom diagrams-a systematic review". *Journal of Orthopaedic and Sports Physical Therapy*, vol. 50, no. 11, pp. 622-631, 2020.

[8] J. Hong, X. Wang, J. Xue, J. Li, M. Zhang and W. Mao. "Clinical characteristics and treatment of adult idiopathic carpal tunnel syndrome accompanied with trigger digit". *Computational and Mathematical Methods in Medicine*, vol. 2022, p. 8104345, 2022.

[9] K. Möllestam, M. Englund and I. Atroshi. "Association of clinically relevant carpal tunnel syndrome with type of work and level of education: A general-population study". *Scientific Reports*, vol. 11, no. 1, p. 19850, 2021.

[10] B. Feng, K. Chen, X. Zhu, W. Y. Ip, L. L. Andersen, P. Page and Y. Wang. "Prevalence and risk factors of self-reported wrist and hand symptoms and clinically confirmed carpal tunnel syndrome among office workers in China: A cross-sectional study". *BMC Public Health*, vol. 21, no. 1, p. 57, 2021.

[11] S. Kumar and M. Muralidhar. "Analysis for prevalence of carpal tunnel syndrome in shocker manufacturing workers". *Advances in Production Engineering and Management*, vol. 11, no. 2, pp. 126-140, 2016.

[12] H. C. Tang, Y. Y. Cheng and H. R. Guo. "Association between hormone replacement therapy and carpal tunnel syndrome: A nationwide population-based study". *BMJ Open*, vol. 12, no. 1, p. e055139, 2022.

[13] B. Evanoff, B. T. Gardner, J. R. Strickland, S. Buckner-Petty, A. Franzblau and A. M. Dale. "Long-term symptomatic, functional, and work outcomes of carpal tunnel syndrome among construction workers". *American Journal of Industrial Medicine*, vol. 59, no. 5, pp. 357-368, 2016.

[14] F. Sharief, J. Kanmani and S. Kumar. "Risk factors, symptom severity and functional status among patients with carpel tunnel syndrome". *Neurology India*, vol. 66, no. 3, pp. 743-746, 2018.

[15] Y. H. Shin, J. O. Yoon, Y. K. Kim and J. K. Kim. "Psychological status is associated with symptom severity in patients with carpal tunnel syndrome". The *Journal of Hand Surgery*, vol. 43, no. 5, pp. 484.e1-484.e8, 2018.

[16] S. Eslami, B. Fadaei, M. Baniasadi and P. Yavari. "Clinical presentation of carpal tunnel syndrome with different severity: A cross sectional study". *American Journal of Clinical and Experimental Immunology*, vol. 8, no. 4, pp. 32-36, 2019.

ORIGINAL RESEARCH ARTICLE UHD JOURNAL OF SCIENCE AND TECHNOLOGY

An Effective Computer-aided diagnosis Technique for Alzheimer's Disease Classification using U-net-based Deep Learning

Fawzi Abdul Azeez Salih[1], Shaniar Tahir Mohammed[2], Tofiq Ahmed Tofiq[2], Hataw Jalal Mohammed[3]

[1]Department of Computer Science, College of Basic Education, University of Sulaimani, Sulaimani, Iraq, [2]Department of Computer Science, College of Science, University of Sulaimani, Sulaimani, Iraq, [3]Charmo Center for Research, Training, and Consultancy, University of Charmo, Sulaimani, Iraq

ABSTRACT

The diagnosis of Alzheimer's disease (AD), a common neurodegenerative disease that impairs thinking and memory abilities in older adults and ultimately results in cognitive impairment and dementia, is made possible in large part by computer-aided diagnosis (CAD). The idea has been to use either machine learning models or deep learning models to develop classification techniques for this disease. CAD techniques and mechanisms have emerged to help and facilitate early detection of this disease as a fundamental step in its treatment plan. As part of our approach, we proposed a model that included the following two pre-processing steps: Contrast Limited Adaptive Histogram Equalization (CLAHE) was utilized to enhance image contrast, especially in low-contrast areas. Normalization was then incorporated to ensure reliable training and faster convergence. A Gray-level co-occurrence matrix technique was used to extract seven texture features from the images following pre-processing: contrast, homogeneity, energy, correlation, variance, dissimilarity, and entropy. After that, these characteristics were added to the model output before the last classification layer. The best hybrid framework out of the five models we examined in this paper was utilized to build a convolutional neural network that can be used to identify AD characteristics from magnetic resonance images. As discussed in Section IV of this article, the U-Net model was selected because of its superior performance. The experimental results demonstrate that this technique showed great accuracy in segmentation and classification for each of the five AD Neuroimaging Initiative categories when a specific diagnosis was made. These results are as follows: Overall, the five classes' final average scores for the four measures were as follows: 94.46% for Accuracy, 94.32% for Precision, 94.49% for Recall, and 94.41% for F1-score.

Index Terms: Alzheimer Diseases, CLAHE, U-Net, Convolutional Neural Network, Magnetic Resonance Imaging, Alzheimer's Disease Neuroimaging Initiative

1. INTRODUCTION

Today, due to the advancements in the fields of biomedical engineering, data acquisition techniques, and data analytics, Computer-aided diagnosis (CAD) systems are used across almost all the fields of medicine [1], and one of the prevalent diseases in medicine fields is the progressive neurodegenerative disorder brain atrophy Alzheimer's disease (AD) [2], which is characterized as a most common neurological disorder that ultimately triggers an irreversible decline in cognitive function sciences [3], because it is a very multifaceted ailment that reasons brain disappointment, then ultimately, dementia ensues. It is a global health problem. (99%) of clinical trials have failed to limit the progression

Access this article online

DOI: 10.21928/uhdjst.v9n1y2025.pp34-43

E-ISSN: 2521-4217

P-ISSN: 2521-4209

Corresponding author's e-mail: Fawzi Abdul Azeez Salih, Department of Computer Science, College of Basic Education, University of Sulaimani, Sulaimani, Iraq. E-mail: fawzi.barznji@univsul.edu.iq

Received: 10-01-2025 Accepted: 09-02-2025 Published: 25-02-2025

of it [1]. The calculated annual fee of dementia is predicted to be a trillion US dollars and is predicted to double by 2030 [4]. According to the World Health Organization (WHO) report, more than 55 million people suffer from dementia worldwide, and more than (60%) of them live in low- and middle-income and learning countries. Every year, there are approximately 10 million new cases [5], and by 2050, it is expected to reach 13.8 million [4]. This indicates that the prevalence of this disease will increase by more than (200%) over the next 15 years [6].

As illustrated in Fig. 1, AD has five stages: (1) AD dementia with severe symptoms, then (2) Late Mild Cognitive Impairment (LMCI), (3) Mild Cognitive Impairment (MCI): which is a condition that precedes dementia but does not meet the criteria for a diagnosis of AD, (4) Early Mild Cognitive Impairment (EMCI), (5) Cognitively Normal (CN): pre-clinical dementia, which is classified by the symptom-free period that occurs between the initial brain lesions and the onset of the first symptoms [7].

The indications of AD typically evolve slowly and gradually, and also patients may show various symptoms at cognitive and behavioral levels; therefore, it can be difficult and complex to diagnose AD. Within this framework, developing innovative diagnostic tools to help diagnosing the disease at an earlier stage is a challenging task. In this context, there has been growing interest in using CAD systems for automatic detection of AD [8]. A variety of CAD approaches have been proposed for the early diagnosis of various stages of AD using Magnetic resonance imaging (MRI) [9], recently provided a non-invasive imaging approach that can detect subtle morphological changes in the brain [3]. MRI-based atrophy measurements are considered valid markers of disease state and progression since atrophy seems to be an inevitable and intrinsic factor of progressive neurodegeneration. Moreover, changes in structural measures, such as ventricular enlargement, entorhinal cortex, whole brain, and temporal lobe volumes, can be associated with modifications in cognitive performance [7]. MRI scans provide detailed insights into blood circulation and cerebral processes. Still, they cannot detect brainwave activity or facilitate communication between tumor cells [8], in addition, dissimilar X-rays, MRI does not emit ionizing radiation, so it can be considered a valuable opportunity to track the development of AD and monitor the effectiveness of treatment [10]. With the development of artificial intelligence (AI) and the great progress in the field of computer vision and deep learning (DL) over the past years, CAD applications in the medical field have become widespread and play an important role in diagnosing diseases, including the subject of our research (AD) [11]. In CAD systems, DL is now making great strides in medical image analysis [12]. DL has increased importance in medical image analysis, driving the pursuit of AI in medical imaging, which is widely accepted for pattern recognition, primarily due to their unique feature of being trainable as a complete program [8]. A huge number of articles and researches have been published through the internet about attendance and the importance of DL in medical images, these ideas and approaches have included individually or mixed (characterization, detection, segmentation, registration, and classification). In the field of medical imaging, especially in the analysis of AD, there is a well-known trend of merging DL models with node segmentation models include several network architectures, such as convolutional neural network (CNN), which are the furthermost used, VGG16, ResNet, U-Net, Mask R-CNN, etc. [13]. The human brain has a structure with many unique features that can be extracted by different CNN models [14]. U-net is one of the greatest prevalent network constructions used typically for segmentation [15], because it is a semantic segmentation network that is constructed on the full CNN, and was sophisticated in 2015 for the processing of medical images [16]. Instead of using single pixels to diagnose disorders, such as Alzheimer's, trends across regions of interest must be analyzed. Pixel-level image quality should be strong, with enough contrast and

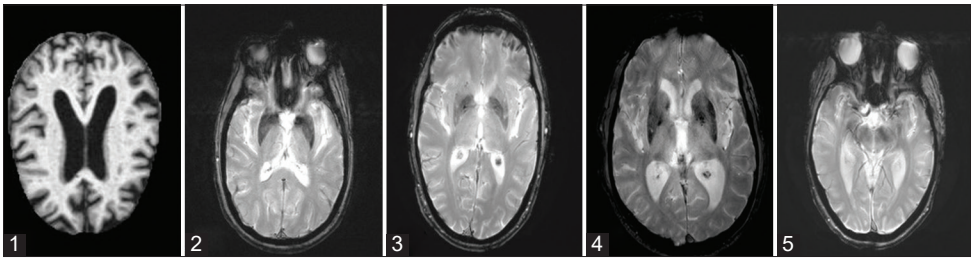


Fig. 1. Samples of magnetic resonance imaging images representing different Alzheimer's disease (AD) stages. (1) AD, (2) Late mild cognitive impairment; (3) Mild cognitive impairment, (4) Early mild cognitive impairment, (5) Cognitively normal.

spatial resolution to reliably identify diseased or anatomical characteristics. The study's minimal pixel quality should be in line with the requirements of the imaging modality. Pixels of poor quality may produce noise or artifacts that jeopardize the diagnostic results and model reliability. Along with advancing quality control measures, these permits employing the measurable features of impairment (area, direction, etc.). This also makes it probable to well understand the tincture of impairment and develop phases to disregard it. To realize this, the U-Net neural network offers for the semantic segmentation of images, where each image pixel is classified as belonging to one of the damaged classes, or to the undamaged part.

2. RELATED WORK

In this part of our article, we will try to provide a comprehensive review of the articles conducted during the past 5 years that is similar to the same topic of our study. Xia *et al.* 2020 [17], proposed a new combined CNN framework for AD detection, and mutually 3D CNN and 3D convolutional long short-term memory (3D CLSTM) were used. They exploit a 3D CNN consist of 6 layers to learn instructive features first, then 3D CLSTM is increased to additional extract the channel-wise higher-level information. The model applied on AD Neuroimaging Initiative (ADNI) dataset and achieved (94.19%) of accuracy rate. Murugan *et al.*, in 2021 [18], employed a DEMentia NETwork (DEMNET) with CNN to extract the discriminative features contained (4) core phases: pre-processing data, balancing dataset consuming Synthetic Minority Over-sampling Technique (SMOTE), Splitting dataset, and classification using DEMNET to detect the dementia stages from MRI obtained from Kaggle using the ADNI dataset to predict AD classes. The proposed DEMNET achieved an accuracy of (95.23%) and an area under curve of (97%). Zhu *et al.* also in 2021 [19], proposed a Patch-Net to generate local illustrations from the brain MRIs. They developed an attention-based pooling block for features mixture and completely -associated layers assisted for final calculations. Their model investigated on ADNI dataset, and they obtained the best result was (92.40%) of accuracy rate. Furthermore, in 2021, Shoaip *et al.* [20], used the ADNI dataset and aimed to propose an interpretable approach to detect AD based on AD diagnosis ontology and the expression of semantic web rule language, by applying an ontology-based method that employs (3) diverse machine learning algorithms, such as random forest, JRip, and J48, after excluding features with a high percentage of missing data, such as DIGITSCOR, AV45, ABETA, TAU, and EcogSPTotal. The proposed classifiers achieved an accuracy of (94.1%) and a precision of (94.3%). Helal *et al.* in 2022, used the ADNI Medical Image dataset and proposed a main objective framework with DL-AD (DL-AHS) based on the U-Net architecture and estimated using the Processing, Analysis, and Visualization technique. They anticipated two architectures for left and right HC segmentation from other brain sub-regions. First utilized simple hyperparameters tuning in the U-Net (SHPT-Net) and the second employed a transfer learning technique in which the ResNet blocks are used in the U-Net (RESU-Net). The result achieved a performance (94.34%) of accuracy rate [21]. In 2023, Noh *et al.* [22], employed spatial and sequential feature extractors, utilized the former U-Net construction in extraction, after that used LSTM to extract temporal features, and executed (4)-step pre-processing to eliminate noise from the fMRI images. In their trained approach, they qualified each of the (3) models by fine-tuning the time measurement. Finally, they revealed an average (96.4%) of accuracy when consuming (5)-fold cross-validation. Furthermore, in 2023, Chen *et al.* [23], proposed a model that directly modeled the brain's organizational networks from DTI. They linked the permanent toolkits, Brain Diffuser, and thwarted additional operational connectivity features and disease-related information by investigating differences in structural brain networks across subjects. They achieved an accuracy rate of (87.83%), Precision (87.83%), Recall (92.66), and F1-score (87.83). In the same year, Bhosale *et al.* 2023 [24], used a U-Net Convolutional Network-based approach to segment AD from ADNI 2D brain MRI images. By implementing a series of convolutional functions using a (3 × 3) filter as the initial design of the U-Net, they used a mixing technique of minimum pooling and average pooling as a hybrid pooling instead of using only maximum pooling. Finally, their updated approach clearly outperformed the original U-Net model, achieving an impressive performance of (91.23%) accuracy. Gupta *et al.* in 2024 [2], conducted an organized evaluation to investigate the estimation of AD on existing toolkits in the ADNI dataset using the Preferred Reporting Item for Systematic Review and Meta-Analysis strategies using ADNI dataset and presented AD Detection Network employment, They achieved results: an accuracy rate of (94.33%), Precision (90.4%), Recall (90.3%) and F1-Score (91.2%). Firdos *et al.* in 2024 [25], explored the effectiveness of CNN constructions, such as UNet, LeNet, and GoogLeNet, and revealed that the CNN model achieved the highest accuracy, with LeNet achieved an accuracy of (97%), UNet at (94%), and GoogLeNet at (51%), using ADNI dataset images. These focus attention on the potential of DL to improve the detection and classification of AD and prepare early

TABLE 1: Summary of related works (They all used the ADNI dataset)

Related work	Preprocessing	Model training	Feature extraction	Classification	Results
Xia *et al*. 2020 [17]	Not detailed	3D CLSTM	Spatial information from the 3D	3D CNN and 3D CLSTM	Accuracy 94.19%
Murugan *et al*., [18]	Spatial rotation, flipping, scaling,	DEMNET	Spatial features.	softmax activation function	Accuracy 95.23% AUC 97%.
Zhu *et al*. [19]	Skull stripping, spatial normalization, and intensity normalization	Dual Attention Multi-Instance Deep Learning (DA-MIDL)	Deep features from each 3D patch	Fully connected layer followed by a softmax activation function	Accuracy 92.40%
Shoaip *et al*. [20]	Skull stripping, intensity normalization, and spatial alignment	Semantic rule-based framework.	Brain volume metrics (e.g., hippocampus, entorhinal cortex)	Semantic reasoning engine	Accuracy 94.1% and precision 94.3%
Helal *et al*. [21]	Skull stripping, spatial normalization, and intensity normalization	Neural Network: U-Net-like architecture	Hippocampus and cortical regions	InceptionV3-TL	Accuracy 94.34%
Noh *et al*. [22]	Motion correction, Slice timing correction, Spatial normalization, Spatial smoothing. Region of Interest (ROI)	Support Vector Machine (SVM)	3D-CNN used to extract spatial features and the rs-fMRI for temporal	Utilization of SVM to classify subjects into respective categories	Accuracy 96.4%
Chen *et al*. [23]	Noise reduction and tensor reconstruction. Spatial alignment. Extraction of fractional anisotropy (FA) and mean diffusivity (MD).	A novel diffusion-based generative model called "Brain Diffuser"	Feature Extraction Net (FENet) to extract the structural attributes from DTI.	Neural network models integrated into the pipeline for disease classification	Accuracy 87.83% Precision 87.83% Recall 92.66% F1-score 87.83%
Bhosale *et al*. [24]	Noise reduction, Normalization, Skull stripping, Resizing.	Enhanced U-Net architecture	Gray matter, white matter, cerebrospinal fluid, Volumetric features.	A shallow neural network trained on extracted features.	Accuracy 91.23%
Gupta *et al*. [2]	Image Registration, Normalization, Skull Stripping	Adversarial Network-based architecture designed for multimodal data	sMRI: gray matter volume, cortical thickness, and hippocampal shape. fMRI: Captured functional connectivity patterns and brain activity networks.	Dense neural network	Accuracy 94.33% Precision 90.4% Recall 90.3% F1-score 91.2%
Firdos *et al*. in 2024 [25]	Resizing: dimensions required by GoogLeNet, LeNet, and UNet. Normalization. Data Augmentation: rotations, flips, and brightness. Segmentation (UNet-specific).	GoogLeNet LeNet UNet	GoogLeNet: multi-scale spatial features. LeNet: Captured basic structural patterns. UNet: Focused on segmenting brain regions.	Effectiveness of CNN constructions, such as UNet, LeNet, and GoogLeNet	Accuracy GoogLeNet 51% LeNet 97% UNet at 94%

AUC: Area under curve, CNN: Convolutional neural network, ADNI: Alzheimer's disease Neuroimaging Initiative

interferences and individual care effortless. The promising results from the CNN model's highpoint, their ability to convert the clinical technique to Alzheimer's, highlighting the importance of technical developments in addressing this incapacitating state. Table 1 below summarizes the related works, and describes (Dataset, Pre-processing, model training, feature extraction, classification, and results).

3. PROPOSED METHOD

In this study, we proposed a hybrid image classification approach (U-Net CNN) to classify the five pre-determined classes of AD. The method was divided into several stages: image pre-processing, feature extraction, and DL-based classifications as shown in Fig. 2 below.

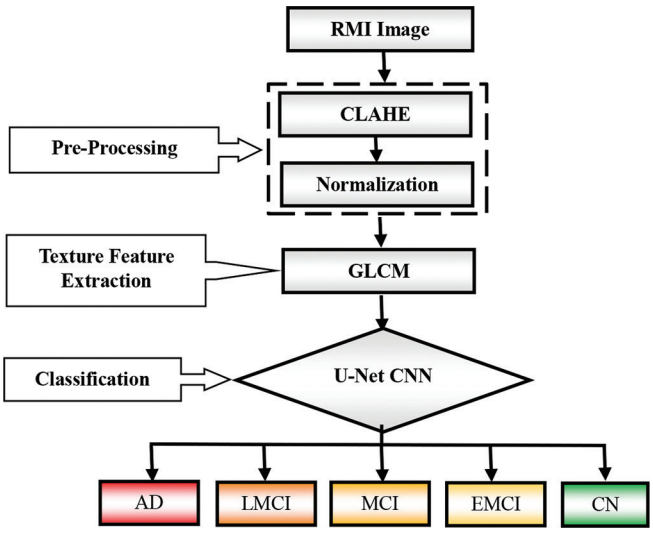


Fig. 2. The block diagram of the proposed approach.

3.1. Pre-Processing

To reduce the (time requested and the learning difficulty) of the proposed model, we increased the contrast level of the images and then we normalized them which is very necessary for detecting and classification of AD cases. This pre-processing stage included two steps:

a. Using contrast limited adaptive histogram equalization (CLAHE) which performed to enhance the contrast of specific ranges by adjusting the intensity levels according to local histograms [26], as shown in Fig. 3. This leads to additional detailed illustration of the crucial structural features of and improve our technique.

b. There are several types of normalization, such as intensity normalization (IN), spatial normalization, Z-score normalization, and numerical normalization. That can be used to remove some variations in the data, such as different subject pose or differences in image contrast, to simplify the detection of subtle differences [21]. In our proposed model, IN was used, where the pixel intensity values of the images are normalized to the range [0, 1] by dividing the pixel values by 255. This kind of normalization confirms that the input data has steady intensity levels, refining convergence throughout training and making the model fewer sensitive to variations in input brightness. Fig. 4 illustrates the result of normalizing on the same images that used in Fig. 3.

3.2. Texture Feature Extraction

Enables the extraction of valuable information for tasks, such as texture classification and segmentation [27]. Everywhere, when the Gray-level co-occurrence matrix (GLCM) is computed, there are seven numerous statistical measures (Contrast, Homogeneity, Energy, Correlation, Variance, Dissimilarity, and Entropy) can be derived from it to characterize the texture and structure of the image. Experimenting with different parameters and features allows for fine-tuning the analysis to specific application requirements, making GLCM a versatile tool in the field of image processing and computer vision [27].

3.3. U-net-based DL Framework

Among various network models, U-Net stands out as the most widely used encoder–decoder model for medical image segmentation [28]. U-net is a neural network model that is usually used for medical image segmentation and its performance has become the baseline for most medical image semantic segmentation tasks. Fig. 5 below demonstrates the universal construction of a basic U-Net neural network [29]. It is proportion and contains two main portions: the compression (i.e. the encoder: left), and the expansion (i.e. the decoder: right). The compression part is a typical CNN structure, contain recurrent convolutions with a (3 × 3) kernel, followed by rectified linear unit (ReLU) operations and max pooling, and with each down-sampling procedure, there is a doubling of feature maps. At the end of each up-sampling, convolution is applied using a (3×3) kernel and a ReLU activation function. As a result of the up-sampling, new pixels are inserted between the existing ones, until the image reaches the wanted size. The final layer uses a 1×1 convolution, which schemes each feature vector onto the anticipated number of classes.

Thus, a Hybrid Framework of a U-net-based CNN style model is proposed for the diagnosis of AD. The core features of the U-net neural network comprise skip connections and a U-shaped structure with symmetrical encoders and decoders. The U-net executes down-sampling operations through the encoder to gain high-level semantic features and up-sampling operations through the decoder to correspondingly restore the high-level semantic feature map to the original image determination. At the same time, the network structure mixes the improved feature with the low-level features through skip connections, which helps the model not only learn the semantic features from MRI scans but also motivates the model to pay attention to the original subtle features. The classification of the images is performed using a hybrid approach, where texture features are learned by a custom U-Net-based CNN. The proposed hybrid framework U-Net-based CNN is applied based on the default U-Net architecture as shown in Fig. 6. It consists of ten layers:

1) Input Layer: Accepts pre-processed image data and texture features (GLCM) as inputs.
2) Down-sampling (encoding layer) involved 2 blocks,

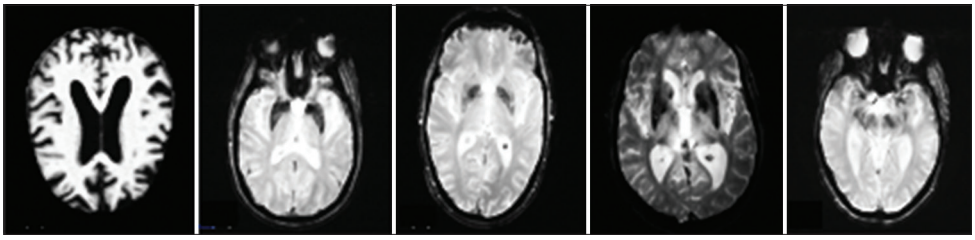


Fig. 3. The effect of applying CLAHE on the images and raising the contrast value.

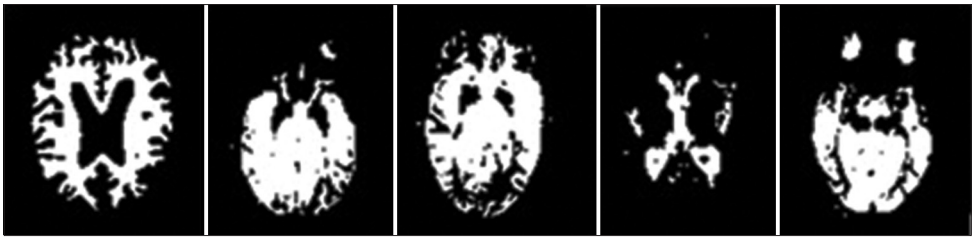


Fig. 4. Normalizing images to the range [0,1].

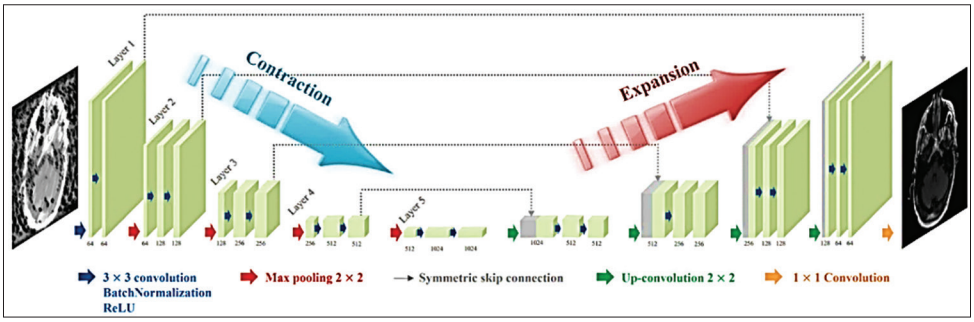


Fig. 5. U-shaped structure of the U-Net neural network [30].

and each block: Applied (conv2D (32) filters, kernel size = (3,3), activation= "ReLU", padding = "same") to reduce feature map dimensions by half and use Max-Pooling 2D (Pool size (2, 2), to reduces spatial dimensions.

3) Bottleneck layer to compress features to the smallest spatial representation Applied (conv2D: High-dimensional (128) filters and Dropout as a Regularization to reduce overfitting.
4) Up-sampling (decoding layer) with Size (2, 2) involved 2 blocks, and each block: Applied (Conv2D: 64 filters, kernel size (3, 3), padding = "same", Skip Connection: Concatenate with corresponding encoder block, Conv2D: 64 filters, kernel size (3, 3), padding = "same")
5) Fusion with GLCM features (Process GLCM texture features using a Dense layer (64 neurons), then combining processed GLCM features with the decoder output through concatenation.
6) Output layer, applied (conv2D layer with filters = number of classes and activation= "softmax" to produce class probabilities.
7) Compilation (loss function and optimizer) layer, involved: (a) Loss Function: categorical cross-entropy, multi-class classification, (b) Optimizer: Adaptive Moment Estimation optimizer, learning rate=$1e^{-4}$, (c) Metrics: Accuracy, Precision, Recall, and F1-score for performance evaluation.
8) Hybrid framework layer: combining spatial (U-Net-based) and texture (GLCM-based) features for classification, the objective was to leverage both spatial features (via CNN) and handcrafted descriptors (via GLCM) and the benefit was to improves classification by capturing complementary patterns, enhancing Alzheimer's detection accuracy.
9) Training the CNN: Pre-processed images and GLCM features are passed through the model, and Labels are one-hot encoded for multi-class support, then Train using augmented data (Techniques include rotation, shifting, and flipping) to improve generalization.
10) Optimization method: used Adaptive Moment Estimation

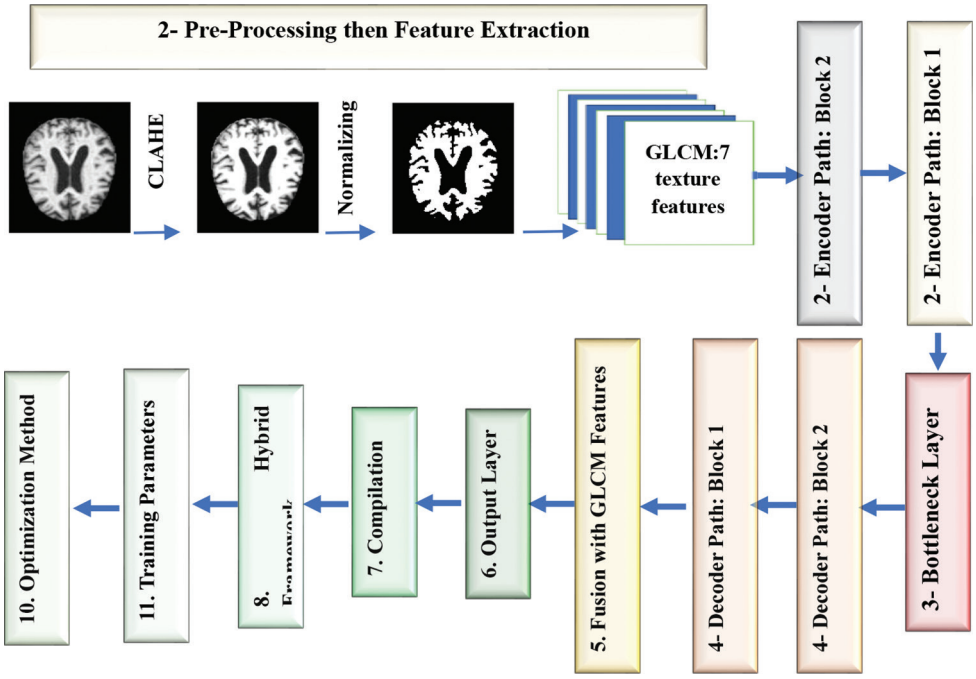


Fig. 6. The proposed hybrid framework U-Net-based convolutional neural network architecture.

optimizer (Adam) optimizer with adaptive learning rates to ensure smooth convergence during training.

11) Training parameters: Define epochs (20) and batch size (defined implicitly by the fit method) to balance speed and convergence.

3.4. Training Process for the Hybrid U-Net and CNN Model

The objective was to train the Hybrid U-Net with GLCM Features to learn spatial, texture-based, and semantic features for AD detection.

3.4.1. Inputs during training

a) Input Images: Pre-processed MRI scans fed into the U-Net encoder.
b) GLCM Features: Extracted texture features (contrast, homogeneity, energy, and correlation) concatenated at the decoder stage.
c) Labels: Ground truth labels (e.g., AD stages or healthy controls), either for segmentation maps (if using U-Net for segmentation) or for classification.

4. EXPERIMENTAL RESULT

4.1. Dataset

A Novel Image Casting and Fusion for Identifying deep Information utilized in this paper was gained from the ADNI information base (www.kaggle.com/datasets/kaushalsethia/alzheimers-adni) which is a comprehensive and widely used collection of MRI images format. ADNI inspires a direction for scientific researchers to main robust investigation and offer feasible evidence with different predictors around the world. The dataset contains a total of 18775 imaging sessions in which the patients or individuals are categorized into two groups (testing and Training) as shown in Table 2 below, and each group had alienated into five classes that are: AD, LMCI, MCI, EMCI, and CN.

TABLE 2: Total 18775 MRI images in ADNI classified into different AD categories

Groups	Classes	Testing images	Training images
Demented	AD	810	7536
	LMCI	72	72
	MCI	233	922
	EMCI	240	240
Non-Demented	CN	1220	7430
Total=18775		2575	16200

LMCI: Late mild cognitive impairment, MCI: Mild cognitive impairment, EMCI: Early mild cognitive impairment, CN: Cognitively normal, AD: Alzheimer's disease, MRI: Magnetic resonance imaging

4.2. Performance Metrics

a) Accuracy is the most common measure used to answer the question "Of all the predictions we made, how many were correct?," therefore ACC is the number of accurate predictions to the whole quantity of predictions. And calculated by Equation (1).

TABLE 3: CNN results without pre-processing

Groups	Classes	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
Demented	AD	92.21	92.45	92.52	92.48
	LMCI	89.48	89.25	89.35	89.30
	MCI	90.41	90.74	91.74	91.24
	EMCI	90.85	90.35	91.40	90.87
Non-Demented	CN	93.98	93.48	92.11	92.79
Average		91.39	91.25	91.42	91.34

LMCI: Late mild cognitive impairment, MCI: Mild cognitive impairment, EMCI: Early mild cognitive impairment, CN: Cognitively normal, AD: Alzheimer's disease, CNN: Convolutional neural network

TABLE 4: CNN results after pre-processing

Groups	Classes	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
Demented	AD	94.06	93.62	94.19	93.90
	LMCI	90.33	89.50	90.06	89.78
	MCI	90.66	90.59	91.84	91.21
	EMCI	91.80	91.51	92.15	91.83
Non-Demented	CN	93.83	93.33	91.96	92.64
Average		92.14	91.71	92.04	91.87

LMCI: Late mild cognitive impairment, MCI: Mild cognitive impairment, EMCI: Early mild cognitive impairment, CN: Cognitively normal, AD: Alzheimer's disease, CNN: Convolutional neural network

TABLE 5: U-Net model's performance after pre-processing

Groups	Classes	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
Demented	AD	95.07	94.32	94.19	94.25
	LMCI	91.34	90.52	90.82	90.67
	MCI	91.58	92.61	92.06	92.33
	EMCI	93.55	92.12	92.28	92.20
Non-Demented	CN	93.83	93.33	91.96	92.64
Average		93.07	92.58	92.26	92.42

LMCI: Late mild cognitive impairment, MCI: Mild cognitive impairment, EMCI: Early mild cognitive impairment, CN: Cognitively normal, AD: Alzheimer's disease

TABLE 6: The final results from the five categories in the ADNI dataset using U-netbased CNN

Groups	Classes	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
Demented	AD	96.08	96.32	96.39	96.35
	LMCI	92.35	92.12	92.22	92.17
	MCI	93.28	93.61	94.61	94.11
	EMCI	93.72	93.22	94.27	93.74
Non-Demented	CN	96.85	96.35	94.98	95.66
Average		94.46	94.32	94.49	94.41

LMCI: Late mild cognitive impairment, MCI: Mild cognitive impairment, EMCI: Early mild cognitive impairment, CN: Cognitively normal, AD: Alzheimer's disease, ADNI: Alzheimer's disease Neuroimaging Initiative, CNN: Convolutional neural network

TABLE 7: Comparative six models experimental results

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
FCN	93.71	93.52	94.30	93.91
SegNet	94.06	94.23	94.51	94.37
ResNet	93.41	94.12	93.97	94.04
DenseNet	94.14	94.09	93.66	93.87
U-Net	94.46	94.32	94.49	94.41%

$$Accuracy = \frac{TP + FP + TN + FN}{TP + FP + TN + FN} \times 100 \quad (1)$$

b) Precision is a metric that gives you the number of true positives to the number of total positives that the model expects. Or we can say "the obtainable of all the positive predictions we completed, how many were correct?", it is calculated by Equation (2).

TABLE 8: Accuracy rate of other approaches with the same ADNI dataset

Techniques	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
Xia *et al*. 2020 [17]	94.19	-	-	-
Zhu *et al*. 2021 [19]	92.40	-	-	-
Shoaip *et al*. 2021 [20]	94.10	94.3	-	-
Helal *et al*. in 2022 [21]	94.34	-	-	-
Chen *et al*. 2023 [23]	87.83	87.83	92.66	87.83
Bhosale *et al*. 2023 [24]	91.23	-	-	-
Gupta *et al*. 2024 [2]	94.33	90.4	90.3	91.2
Firdos *et al*. in 2024 [25]	94			
Proposed Approach	94.46	94.32	94.49	94.41

$$Precision = \frac{TP}{TP + FP} \times 100 \quad (2)$$

c) Recall focuses on how good the model is at the outcome of all the positives. It is also entitled "true positive rate" and replies the question "Out of all the data points that should be predicted as true, how many did we acceptably predict as true?", Recall is calculated by Equation (3).

$$\text{Recall} = \frac{TP}{TP + FN} \times 100 \quad (3)$$

d) F1-score: Balances precision and recall, making it most useful when dealing with imbalanced datasets or unequal error costs, F1-score is calculated by Equation (4).

$$\text{Recall} = \frac{2 \times \text{Recall} \times \text{Precision}}{\text{Recall} + \text{Precision}} \times 100 \quad (4)$$

4.3. Performance Evaluation

As a baseline DL method, we first assessed the CNN model without and then after pre-processing, the findings are shown in Tables 3 and 4.

Next, we evaluated the U-Net model's performance independently after pre-processing, as indicated in Table 5.

Finally, the hybrid approach, which combines CNN and U-Net, was then used to examine the possible advantages of this integration. Table 6 displays the final results for the four metrics are detailed with the final average for each metric for the five categories in the ADNI dataset.

4.4. Models Validation

To validate our technique and ensure the segmentation effect of the proposed hybrid U-Net framework model, other four models, including FCN, SegNet, Resnet, and Densenet were tested on the same prepared dataset, the results in Table 7 showed the strength of our proposed model.

4.5. Performance Comparison

More tests were conducted to evaluate our suggested technique's performance by comparing it with other techniques. Our suggested hybrid technique achieves the greatest performance (accuracy, precision, recall, and F1-score), as shown by the findings in Table 8. Some of certain cells are left blank since all of the most recent methods either tested their strategy only in terms of accuracy or combined accuracy with recall or F1-score.

5. CONCLUSION

Even though DL was initially successful in clinical practice, there are still difficulties in identifying complicated lesions and many intersecting diseases, which calls for the development of more DL-based approaches. When it comes to clinical intelligence-guided decision-making, these analytical endeavors include identifying barriers, creating prediction models, and other essential components that form the basis. The effectiveness of the U-net CNN model was demonstrated by obtaining final average results for the four measures: accuracy (94.46%), precision (94.32%), recall (94.49%), and F1-score (94.41%) as overall rates. The experiment results demonstrate that skip connections and deep supervision can improve the classification model's performance. The U-net CNN model was applied to RMI images from the ADNI dataset, which specializes in AD diagnosis.

REFERENCES

[1] A. Bhandarkar, P. Naik, K. Vakkund, S. Junjappanavar, S. Bakare and S. Pattar. "Deep learning based computer aided diagnosis of Alzheimer's disease: A snapshot of last 5 years, gaps, and future directions". Springer, Berlin, pp. 1-62, 2024.

[2] M. Gupta, R. Kumar and A. Abraham. "Adversarial network-based classification for Alzheimer's disease using multimodal brain images: A critical analysis". *IEEE Access*, Vol. 12, pp. 48366-48378, 2024.

[3] S. Mu, S. Shan, L. Li, S. Jing, R. Li, C. Zheng and X. Cui. "DMA-HPCNet: Dual Multi-level attention hybrid pyramid convolution neural network for Alzheimer's disease classification". *IEEE Access*, vol. 32, pp. 1955-1964, 2024.

[4] H. A. Helaly, M. Badawy and A. Y. Haikal. "Deep learning approach for early detection of Alzheimer's disease". *Cognitive Computation*, Vol. 14, pp. 711-1727, 2022.

[5] WHO. "*Dementia*". World Health Organization, 2023. Available from: https://www.who.int/health-topics/dementia#tab=tab_1 [Last accessed on 2024 Dec 02].

[6] A. A. Fakoya and Parkinson S. "A novel image casting and fusion for identifying individuals at risk of alzheimer's disease using MRI and PET Imaging". *IEEE Access*, Vol. 12, pp. 134101-134114, 2024.

[7] J. Silva, B. C. Bispo, P. M. Rodrigues and Alzheimer's Disease Neuroimaging Initiative. "Structural MRI texture analysis for detecting Alzheimer's disease". *Journal of Medical and Biological Engineering*, vol. 43, pp. 227-238, 2023.

[8] T. Mahmood, A. Rehman, T. Saba, L. Nadeem and S. A. O. Bahaj. "Recent advancements and future prospects in active deep learning for medical image segmentation and classification". *IEEE Access*, Vol. 11, pp. 113623-113652, 2023.

[9] S. A. Javid and M. M. Feghhi. "*Early Diagnosis of Alzheimer's Disease from MRI Images with Deep Learning Model*". IEEE, United States, pp. 1-7, 2024.

[10] R. G. Akindele, S. Adebayo, P. S. Kanda and M. Yu. "AlzhiNet: Traversing from 2DCNN to 3DCNN, towards early detection and diagnosis of Alzheimer's disease". *arxiv.org*, Vol. 60, no. 12, pp. 1-13, 2024.

[11] S. Y. Lu. "A short survey on computer-aided diagnosis of Alzheimer's disease: Unsupervised learning, transfer learning, and other machine learning methods". *Scilight, AI Medicine*, Vol. 1, pp. 1-8, 2023.

[12] S. Sadek and Z. F. Makki. "*A Review of AI Techniques Using MRI Brain Images for Alzheimer's Disease Detection*". IEEE, United States, pp. 76-82, 2023.

[13] B. Youssef, A. Alksas, A. Shalaby, A. Mahmoud, E. Bogaert, N. S. Algahmdi, A. Neubacher, S. Contractor, M. Ghazal, A. Elmaghraby and A. El-Baz. "Integrated deep learning and stochastic models for accurate segmentation of lung nodules from computed tomography images: A novel framework". *IEEE Access*, Vol. 11, pp. 99807-99821, 2023.

[14] R. Al-Amri, R. K. Murugesan, M. Man, A. F. Abdulateef, M. A. Al-Sharafi and A. A. Alkahtani. "Anomaly analysis of Alzheimer's disease in PET images using an unsupervised adversarial deep learning model". *Applied Sciences*, vol. 11, no. 5, pp. 1-17, 2021.

[15] R. Yousef, G. Gupta, N. Yousef and M. Khari. "A holistic overview of deep learning approach in medical imaging". *Multimedia Systems*, Vol. 28, pp. 881-914, 2022.

[16] I. Konovalenko, P. Maruschak, J. Brezinová, O. Prentkovskis and J. Brezina. "Research of U-Net-based CNN architectures for metal surface defect detection". *Machines*, Vol. 10, no. 5, pp. 1-19, 2022.

[17] Z. Xia, G. Yue, Y. Xu, C. Feng, M. Yang and T. Wang. "*A novel end-to-end hybrid Network for Alzheimer's disease detection using 3D CNN and 3D CLSTM*". United States: IEEE, pp. 1-13, 2020.

[18] S. Murugan, C. Venkatesan, M. G. Sumithra, X. Z. Gao, B. Elakkiya, M. Akila and S. Manoharan. "Demnet: A deep learning model for early diagnosis of Alzheimer diseases and dementia from MR images". *IEEE Access*, vol. 9, pp. 90319-90329, 2021.

[19] W. Zhu, L. Sun, J. Huang, L. Han and D. Zhang. "*Dual attention multi-instance deep learning for Alzheimer's disease diagnosis with structural MRI*". IEEE, United States, Vol. 40, no. 9, p. 2354-2366, 2021.

[20] N. Shoaip, A. Rezk, S. El-Sappagh, T. Abuhmed, S. Barakat and M. Elmogy. "Alzheimer's disease diagnosis based on a semantic rule-based modeling and reasoning approach". *Computers, Materials and Continua CMC*, Vol. 9, pp. 3531-3548, 2021.

[21] H. A. Helaly, M. Badawy and A. Y. Haikal. "Toward deep MRI segmentation for Alzheimer's disease detection". *Neural Computing and Applications*, Vol. 34, pp. 1047-1063, 2022.

[22] J. H. Noh, J. H. Kim and H. D. Yang. "Classification of Alzheimer's progression using fMRI data". *Sensors* (*Basel*). Vol. 23, no. 14, pp. 1-14, 2023.

[23] X. Chen, B. Lei, C. M. Pun and S. Wang. "*Brain Diffuser: An End-to-end Brain Image to Brain Network Pipeline*". Springer Nature, Germany, pp. 16-26, 2023.

[24] T. Bhosale, M. Gulame, B. Shendkar, P. Kadam, R. More and R. Mali. "Alzheimer's Disease MRI Image Segmentation Based on the Enhanced U-Net". In: *IEEE, International Conference on ICT in Business Industry and Government (ICTBIG)*, pp. 1-5, 2023.

[25] S. M. Firdos, M. Z. Mehack, S. A. Muskan, A. BibiNadeefa and S. Kamepalli. "Enhancing Alzheimer's Disease Prediction Through Deep Learning Models: A Comparative study of GoogLeNet, LeNet, and UNet". In: *First International Conference on Innovations in Communications, Electrical and Computer Engineering (ICICEC)*, IEEE, 2024.

[26] P. K. Pandey, J. Pruthi, S. Alzahrani, A. Verma and B. Zohra. "Enhancing healthcare recommendation: Transfer learning in deep convolutional neural networks for Alzheimer disease detection". *Front* (*Lausanne*), Vol. 11, pp. 1-12, 2024.

[27] S. U. Khan, N. Islam, Z. Jan, K. Haseeb, S. I. A. Shah and M. Hanif. "A machine learning-based approach for the segmentation and classification of malignant cells in breast cytology images using gray level co-occurrence matrix (GLCM) and support vector machine (SVM)". *Neural Computing and Applications*, vol. 34, pp. 8365-8372, 2022.

[28] Y. Yuan and Y. Cheng. "Medical image segmentation with UNet-based multi-scale context fusion". *Scientific Reports*, Vol. 14, pp. 15687, 2024.

[29] S. H. Kang and Y. Lee. "Motion artifact reduction using U-net model with three-dimensional simulation-based datasets for brain magnetic resonance images". *Bioengineering* (*Basel*), Vol. 11, no. 3, pp. 227, 2024.

REVIEW PAPER UHD JOURNAL OF SCIENCE AND TECHNOLOGY

A review: Multi-Objective Algorithm for Community Detection in Complex Social Networks

Mariwan Wahid Ahmed[1], Kamaran Hama Ali. A. Faraj[2]

[1]Department of Computer, College of Science, University of Sulaimani, Sulaymaniyah, Iraq, [2]Department of Computer Engineering, College of Engineering, Knowledge University, Erbil 44001, Iraq

ABSTRACT

Recently, research on multi-objective optimization algorithms for community detection in complex networks has grown considerably. Community detection based on multi-objective algorithms (MOAs) in complex social networks is a fundamental scheduler, and it supports knowing the dynamics of a society, finding influential groups, and improving information dissemination. The traditional methodologies often cannot cope with the features that real-world network usually present, related to optimizing various and sometimes conflicting objectives. This paper provides an overview of some recent works on MOAs for community detection in complex social networks. This paper will explore the balance of the reached objectives, such as modularity, community size, and edge density. Which are analyzed by 15 different approaches in order to choose from works published during the period 2019–2024. These strengths and limitations of various MOAs are reviewed with a comparative analysis to provide insights into both the effectiveness and computational efficiency of these methods. The present trends and future research are discussed that underline the need for the development of solutions to be more adaptive and scalable in coping with the gradually increasing complexity of social networks.

Index Terms: Meta-heuristic, Multi-Objective Algorithm, Community Detection, Complex Networks, Optimization and Objective

1. INTRODUCTION

Optimization is a field that combines computer science and mathematics to develop methods for solving complex optimization problems. To solve these issues, one objective or multi-objectives must be maximized or minimized depending on optimization variables. The optimization variables may be real or integer values [1]. There are many optimization algorithms designed for many purposes, such as computer technology, economics, engineering, medicine, and logistics. The aim of these algorithms is to find the best solution to an optimization problem [2]. There are three types of optimization:

1.1. Single Objective Optimization

This type of optimization is used when there is only one objective. Other essential objectives are ignored or even have an impact on them [3].

1.2. Multi-Objective Optimization

Algorithms with a set of objectives (typically consist of two or three objectives) are named by multi-objective algorithms (MOAs). MOAs try to identify an optimal solution or more than one solution to an optimization problem by maximizing or minimizing these objectives [4].

Access this article online

DOI: 10.21928/uhdjst.v9n1y2025.pp44-54

E-ISSN: 2521-4217

P-ISSN: 2521-4209

Corresponding author's e-mail: Mariwan Wahid Ahmed, Department of Computer, College of Science, University of Sulaimani, Sulaymaniyah, Iraq. E-mail: mariwan.ahmed@univsul.edu.iq

Received: 20-11-2024 Accepted: 09-02-2025 Published: 27-02-2025

1.3. Many-Objective Optimization

Multi-objective optimization involving more than three objectives is known as many-objective optimization algorithms [5].

As mentioned previously, the field of optimization is used for solving many different problems; one of them is community detection in complex social networks. There are many real-world complex systems that can be demonstrated as complex networks, such as technology networks, social networks, web networks, and biological networks [6]. First, we need to understand what the meaning of community is. A community is a set of entities that are more strongly connected to each other than the other entities within the network [7]. These communities need some techniques, such as optimization algorithms, to be detected within the network. The techniques of community detection play a crucial role in understanding the functionality of complex networks [8]. This process is used to find hidden structures of communities in complex networks and it can be used to find the topology structures of complex networks and understand what the functions of complex networks are [9]–[12]. In mathematics, complex networks can be demonstrated as graphs, where nodes in a graph are denoted as vertices in the network and links can be used to show the edges of a network [13]. Fig. 1 shows a simple community detection using graph theory.

Nowadays, many optimization algorithms are proposed to address the issue of community detection, such as greedy algorithms and meta-heuristic algorithms. However, the greedy technique is not performing well for detecting communities in large complex networks [14]. However, meta-heuristic algorithms play a crucial role in detecting communities in complex social networks. There are many different meta-heuristic algorithms. The vast majority of meta-heuristics belong to algorithms inspired by nature, such as genetic algorithms (GAs), particle swarm optimization (PSO), and ant colony optimization [15]. A number of them belong to non-nature-inspired algorithms, for example, Tabu Search [16] and Iterated Local Search [17].

This review paper reviews some state-of-the-art algorithms based on multi-objective optimization for detecting communities in complex social networks. Furthermore, some of these approaches enhanced meta-heuristic optimizations to detect high-quality communities. Moreover, another made is a combination of the meta-heuristic algorithms

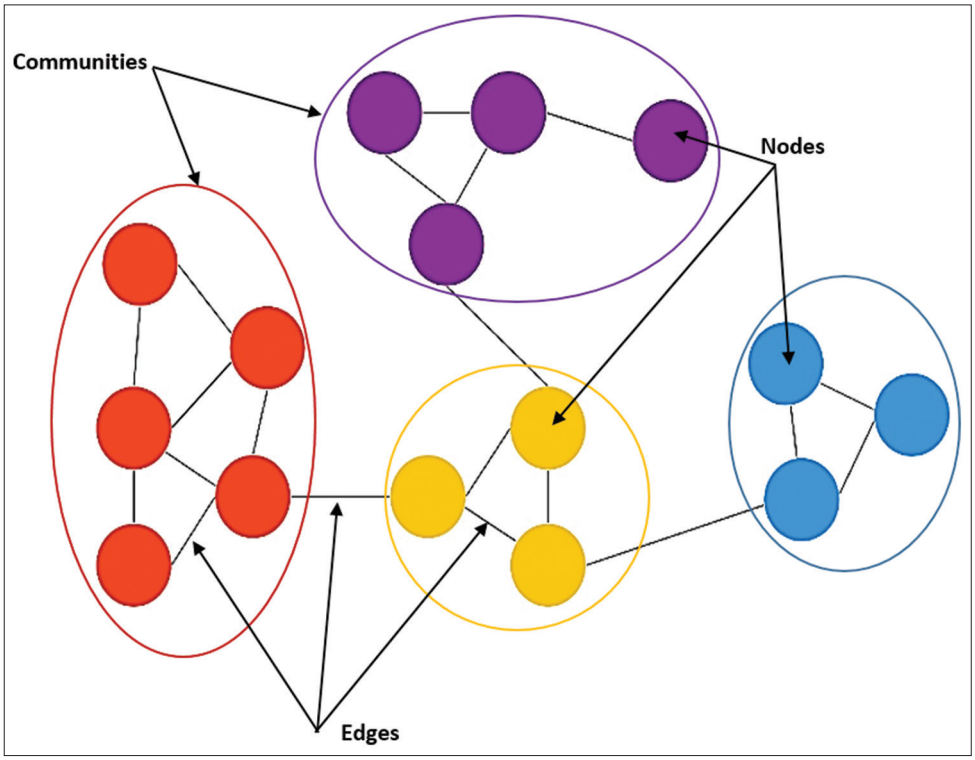


Fig. 1. Example of showing community structure by a graph theory.

to perform better quality of detection. The detail is presented in Section 3.

The rest of the paper is organized as follows: Section 2 is a research strategy; the strategy of the research is explained here. Section 3 presents a literature review; in this section, we discuss the contributions of 15 different researchers. Section 4 is methodology; in this section, the common datasets and evaluation metrics are explained. Section 5 is a comparison and discussion; this section is about the comparison between approaches in terms of objectives, some popular problems of algorithms, solutions, challenges, and future work based on gaps in the reviewed algorithms. The final section is a conclusion; the paper is concluded here.

2. RESEARCH STRATEGY

This paper presents a review of the available literature on MOAs for community detection in a complex social network. Hence, a few different multi-objective strategies that work for community detection in different complex networks are presented. The approaches selected for consideration in supporting the goals of our research have been based on originality and thorough coverage of significant subjects related to MOAs for community detection. In this review paper, fifteen different approaches have been selected from the range of 2019–2024. Moreover, most of the research in our paper is from well-renowned conferences and journals. Furthermore, this literature presents and discusses methodologies and improvements for each selected approach. Table 1 shows the characteristics of fifteen different papers related to community detection in a complex social network by a MOA.

All of these papers are classified into 6 categories. Each category shows the number of papers that were published in the mentioned year that are presented in Fig. 2.

Furthermore, each of those researchers used more than one complex network to test and compare their algorithm with previous algorithms. Table 2 shows the number and names of the networks used by each paper to test the result.

TABLE 1: Presents the characteristics of all 15 papers in our literature

References	Title	Year of publication
[18]	A multi-objective multi-agent optimization algorithm for the community detection problem	2019
[19]	An Enhanced Multi-Objective Evolutionary Algorithm with Decomposition for Signed Community Detection Problem	2020
[20]	A Compression Based Multi-Objective Evolutionary Algorithm for Community Detection in Social Networks	2020
[21]	Evolutionary Multi-Objective Optimization Algorithm for Community Detection in Complex Social Networks	2021
[22]	A Parallel multi-objective evolutionary algorithm for community detection in large-scale complex networks	2021
[23]	Multi-objective NSGA-II-based community detection using dynamical evolution social network	2021
[24]	A fast variable neighborhood search approach for multi-objective community detection	2021
[25]	A Multi-Objective Evolutionary Algorithm with Neighbor Node Centrality for Community Detection in Complex Networks	2022
[26]	A Multi-Objective Evolutionary Algorithm Based on Mixed Encoding for Community Detection	2023
[27]	A Two-Stage Multi-Objective Evolutionary Algorithm for Community Detection in Complex Networks	2023
[28]	A Macro-Micro Population-Based Co-Evolutionary Multi-Objective Algorithm for Community Detection in Complex Networks	2023
[29]	Multi-objective Optimization Overlapping Community Detection Algorithm based on Subgraph Structure	2023
[30]	A Multi-Objective Pigeon-Inspired Optimization Algorithm for Community Detection in Complex Networks	2024
[31]	Two-stage multi-objective evolutionary algorithm for overlapping community discovery	2024
[32]	Community Detection in Social Networks Using a Local Approach based on Node Ranking	2024

NSGA: Non-dominated sorting genetic algorithm

TABLE 2: Presents a number and name of networks used to evaluate the algorithm by each paper in the literature

References	Number of networks used by each paper	Network's formal name
[18]	5	Karate, Les Misérables, Bernard, Grevy's zebra, Facebook
[19]	5	Karate, Dolphins, Football 2000, Football 2001, Krebs
[20]	11	Karate, Dolphins, Football, Polbooks, Co-authors, Email, Netscientist, Facebook, GR_QC, GC_Hep_TH, GC_Hep_PH.
[21]	4	Karate, Dolphin, Football, Books about US Politics.
[22]	10	Football, Net-science, blogs, ca-GrQc, ca-HepTh1, ca-HepTh2, ca-AstroPh, ca-CondMat, loc-Brightkite, loc-Gowalla.
[23]	3	Last.fm, Douban, SYNFIX.
[24]	10	Karate, Dolphins, Football, netscience, jazz, musae_DE_ edgesnetwork, musae_ENGB_ edgesnetwork, musae_ES_ edgesnetwork, musae_FR_ edgesnetwork, musae_RU_ edgesnetwork.
[25]	9	Karate, Dolphin, Football 2000, Football 2001, Kreb's books, SFI, Jazz, Netscience, Power grid.
[26]	4	Karate, Dolphins, Books, Football.
[27]	4	Karate, Dolphin, Football, Polbooks
[28]	14	Karate, Dolphin, football, the Books about US politics, the Yeast PPI dataset, the Blogs network, the PGP network, Ca-GrQc, Ca-HepTh1, Ca-HepTh2, Ca-AstroPh, Ca-CondMat, Epinions, and Enron-large.
[29]	4	Karate, dolphin, football, American political books network.
[30]	3	Karate, Dolphin, Football
[31]	9	Karate, Dolphin, Football, Polbook, Email, Jazz, SFI, Y2H, Yeast-D2.
[32]	18	Karate, Dolphins, PolBooks, Football, SFI, NetScience, Email, PowerGrid, PGP, GrQc, ca-AstroPh, ca-HepTh, ca-HepPh, Condmat-2003, Condmat-2005, Email Enron, Collaboration, Internet.

3. LITERATURE REVIEW

In this section, we discussed the methodology and contribution for each selected paper. The algorithms are categorized into three categories (modified evolutionary algorithms, hybrid evolutionary algorithms, and swarm intelligence algorithms).

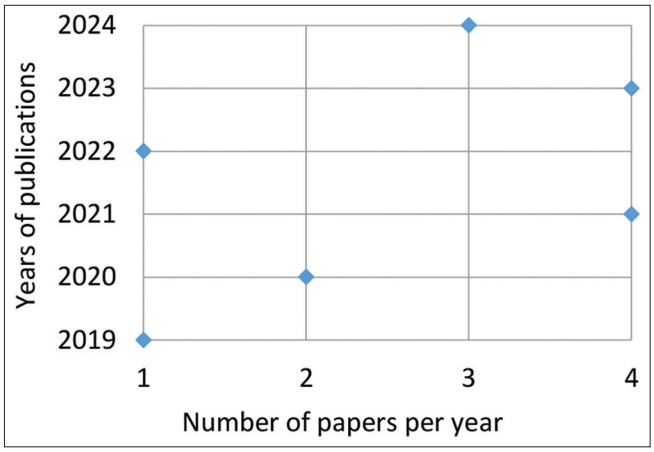


Fig. 2. Demonstrates the number of selected paper per year.

3.1. Modified Evolutionary Algorithms

Hosseinian and Baradaran [18] proposed a new approach named a multi-objective multi-agent optimization algorithm (MAOA) for detecting community problems in social networks. This algorithm optimizes each objective simultaneously to obtain enhanced accuracy and efficiency in the detection of communities in the networks. This algorithm introduces an agent-based multi-objective for detecting communities. The agents are organized into groups of leaders and active agents. By having this, the algorithm used a grouped structure. By enhancing the agent's environment and behaviors, the algorithm effectively facilitated the solution space exploration. Another contribution in this approach is integrating the concept of Pareto dominance to ensure the algorithm can approximate the Pareto optimal front and identify solution efficiency. Experimentally, there are two main contributions: Performance improved by using a multi-agent strategy and providing better initialization to make more applicability in large and more complex networks. Experimental results were tested on some real-world networks and compared with three meta-heuristic algorithms. The result demonstrates that the MAOA performs better in terms of accuracy and efficiency.

Another approach based on a multi-objective evolutionary algorithm with decomposition (MOEA/D) is proposed by Abdulrahman *et al.* [19]. It is decomposing the multi-objective optimization problem into several sub-problems efficiently. The aim of this method is to balance the quality of detecting communities with the distribution of solutions through the Pareto front and find a different set of Pareto

optimal solutions. MOEA/D focused on enhancing mutation operators to improve performance. Unlike the traditional mutation operator, this randomly selects a node for mutation. In this approach, the internal and external connections between nodes have been calculated. Nodes with low internal and high external connections are called positive connections and nodes with high internal and low external connections are called negative connections. The mutation operator has been applied over the positive connections. Having positive and negative connections helps the efficiency of the algorithm. A variety of real-world networks were selected to evaluate the performance of MOEA/D. The result of this approach is outperformed compared with the traditional methods. It provides a better structure for community detection in dynamic networks.

In the same year of [19], Liu *et al.* [20] designed an enhanced algorithm using a multi-objective evolutionary algorithm (MOEA). To optimize the process of community detection, the proposed algorithm performed a compression-based method to optimize two objectives, such as maximizing the modularity and minimizing the number of communities. The compression process was applied over the network topologies to obtain smaller networks. The main purpose of this approach is improving efficiency by compressing networks because the majority of algorithms for community detection have a problem of computational time. Furthermore, to prefer a better community structure, both initialization and mutation processes have been improved. The approach conducts extensive experiments on some datasets and the result outperforms some existing methods in terms of computational efficiency and accuracy. The proposed algorithm has the ability to exploit in the search of an optimal solution and make a balance between exploration and exploitation. The results are tested on some different datasets and compared with some state-of-the-art algorithms. The proposed algorithm obtained higher accuracy and performance.

Another paper by Shaik *et al.* [21] presents a novel algorithm to detect communities in complex social networks using three objectives. Optimizing three objectives is a main novelty because traditional approaches primarily focused on single or two objectives. There are two variants of a non-dominated sorting GA (NSGA) III introduced. The first is NSGA-III-KRM (kernel k-means, ratio cut, and modularity) and the second variant is NSGA-III-CCM (community score, community fitness, and modularity). Moreover, a new measurement of the ranking mechanism for Pareto solutions by using a ratio of hyper-volume to inverted generational distance is produced to enhance the Pareto set evaluation. This research addresses some limitations and improves the performance metrics of previous approaches. To evaluate the result of this algorithm, four network datasets are selected and compared with some existing state-of-the-art methods. The algorithm achieved a good result compared with other algorithms.

A new parallel MOEA (PMOEA) was proposed by Su *et al.* [22]. It is designed to detect communities in large-scale complex networks. The first operation performed by PMOEA is to identify communities based on specific nodes (key nodes) rather than the entire network. After that, it executes multiple copies of a MOEA to detect communities linked to each key node using a parallel mechanism. Another contribution in this approach is enhancing the mutation and crossover operators to capture a better community in the network. Through performing the parallel process, the computational efficiency and quality of detection are enhanced. Experiments of this algorithm indicate that it works better than other evolutionary and non-evolutionary algorithms, showing that it can handle networks with up to 200,000 nodes.

Based on the NSGA II algorithm, the dynamic community detection (DCD) system proposed by Alkhalec Tharwat *et al.* [23]. The proposed algorithm explains the growing need for community detection methods effectively in dynamic social networks. DCD is the main goal of this algorithm, while networks change over time. It utilizes a multi-objective optimization framework by using the specialty of NSGA-II and involves formulating the community detection problems in social networks where the state of networks changes over time. While the network is dynamic, the proposed algorithm identified communities at different time points. This algorithm is able to be used in various applications, such as recommender systems, analysis of social media, and retrieving information. The result of this algorithm obtained better performance compared with classical GAs.

Another approach, using a multi-objective GA (MOGA-Net), is proposed by Abbood *et al.* [25] for community detection in complex networks. The proposed approach uses the NSGA-II for finding globally non-dominated solutions, which guarantees that no other possible partition is superior for both objectives. The proposed algorithm introduced a novel scoring model (intra-score and inter-score). They assist the algorithm to capture better community structure. A new contribution of this approach is identifying the healthy and infected communities during COVID-19. This study compares MOGA-Net's performance with other state-of-

the-art methods using some real networks, showing how efficiently it creates accurate community structures. This algorithm also introduces a new scoring model that quantifies the density of internal connections (intra-score) and the sparsity of inter-community connections (inter-score).

There is an additional contribution by Zhu *et al.* [27] is a fundamental task that aids the comprehension of complex networks. The proposed approach implements a two-stage MOEA that strengthens community detection through optimizing multiple objectives simultaneously. First, in the initial stage, the individual similarity parameters are targeted to find possible communities; then, in the second stage, distinct crossover operators are employed with respect to the characteristics of the detected communities. The algorithm is developed to handle challenges arising in traditional community-detection methods, which face the complexity and dynamics of real-world networks. Furthermore, it enhanced the boundary-independent nodes by applying the second-stage strategy. They further validate their approach through extensive experiments on many datasets, showing that their algorithm performs better than the state-of-the-art approaches in both accuracy and computational efficiency.

In the study of Zhang *et al.* [28], a new algorithm named MMCoMO (Macro-Micro Population-Based Co-Evolutionary Multi-Objective) algorithm was proposed and it's different than the traditional MOEAs. They worked with a single population at the first steps, which lead to a limiting balance between exploration and exploitation. On the other hand, the MMCoMO employs macro-population and micro-population as two types of population. The main contribution in this approach identified the macro and micro populations. The macro-population emphasizes exploration. During this process, the network quickly portioned to detect the approximate community structure. Meanwhile, the micro-population focuses on exploitation to achieve more precise community configurations the structures of the network through local search are refined. In addition to quality enhancement, the MMCoMO improves the computational efficiency of detecting communities compared with existing MOEAs.

After that, Cai *et al.* [31] provide a fast two-stage MOEA for the identification of the overlapping community in networks. The main goal of this algorithm is to discover overlapping communities using a two-stage MOEA. The first stage aims at identifying the high-quality non-overlapping communities by using the population initialization technique using the degree central nodes. It also increases the robustness of community division with respect to the existing methods. In the second stage, the algorithm picks out additional nodes from the networks that have previously been categorized in non-overlapping communities as central nodes. An information feedback model is used to adjust a fuzzy scale for thresholding to improve the identification of overlapping nodes. Another contribution in this approach is identifying a new initialization of the population using central nodes based on node degree for better partitioning. The proposed algorithm's efficiency is tested through experiments on some different networks and compared with other algorithms based on the modularity and accuracy value of community structure and achieved higher results.

Sheykhzadeh *et al.* [32] proposed a novel local community detection algorithm based on node ranking in social networks. It is referred to as LCD-SN local method community detection algorithm in social networks and is designed to overcome the limitations of previous approaches that usually have low accuracy and high computational time. Node ranking means how nodes interact and are ranked relative to community nodes and their connectivity. The algorithm has been designed to look for communities of densely connected nodes and relatively scattered nodes in between communities. First-degree and second-degree neighbor nodes are used by LCD-SN to build the communities, as a result of which accuracy and determinism are improved while seed nodes are not required. The algorithm starts with node scaling and defining the important nodes with the help of local characteristics and then it creates the primary groups with the node and its first-order neighbors. Communities are identified in the last step in the process known as post-processing. The paper also presents a new measure for ranking the nodes in the network, which makes community detection even better. Experimental analysis shows that LCD-SN is useful for finding communities with a flexible approach to the trade-off between time and solution quality.

3.2. Hybrid Evolutionary Algorithms

Pérez-Peló *et al.* [24] present a hybrid meta-heuristic approach for detecting communities in networks by combining variable neighborhood search (VNS) and Greedy randomized adaptive search procedure (GRASP). The main contribution in this approach is improving the efficiency of search by combining VNS and GRASP. Moreover, the limitations of traditional community detection algorithms are highlighted, especially these challenges that are associated with high modularity that suffers from the limitation of resolution. Based on bi-objective community detection problems, this algorithm aims to optimize multiple objectives simultaneously, which leads to a reliable community structure. As a result, the algorithm shows effectiveness compared with the existing approaches and is good for these applications that are used for analyzing complex networks.

In this direction, Yang *et al.* [26] have proposed a hybrid approach to community detection in complex networks. The main contribution in this algorithm is identifying a mixed encoding strategy by combining locus-based and label-based representation. Having this combination helps the algorithm to represent a valid community structure and is more flexible. Furthermore, this study has recognized the difficulties that were posed by methods based on traditional modularity maximization, especially with regard to network topology issues and invalid solution generation. To improve these issues, this paper proposes a MOEA based on the integration of label-based and locus-based representations. The experimental results have shown that this approach outperforms existing algorithms concerning effectiveness and efficiency. These results confirm that this algorithm is a competitive approach when compared with traditional methods of community detection and consequently.

3.3. Swarm Intelligence Algorithms

As a part of swarm intelligence, Li [29], proposed a new overlapping community detection algorithm based on subgraph structures and multiple objective optimization techniques. The main contribution in this approach is transforming overlapping nodes into clique nodes. Due to the possible overlap in real-world communities, the algorithm herein proposed employs the k-core decomposition to find maximum cliques, which form the building blocks for the weighted graph. Instead of randomly initializing the population, the proposed algorithm leverages the k-core decomposition for better initial community partitioning. The algorithm integrates a PSO to improve search accuracy and convergence speed. The result of this algorithm compared with similar community detection algorithms on some real-world networks and demonstrated a decent performance.

Then, Yu *et al.* [30] introduced a new algorithm named multi-objective pigeon-inspired optimization (MOPIO). This algorithm performed three main steps: Initialization, search, and mutation methods. It starts by constructing a solution representation of the community structure to evaluate two objective functions to assess community quality. The main contribution of this research is integrating pigeon-inspired optimization for community detection. It enhanced the applicability of the algorithm for complex networks. Furthermore, to address the problem of misclassification of boundary nodes, they proposed a novel strategy through the mutation process. The result of this work is evaluated based on different networks and it performs better for detecting community structures compared with other existing approaches. This algorithm not only enhanced the accuracy of community detection but also offered a framework compactable with various networks.

4. METHODOLOGY

This review paper explains the methodologies and contributions of fifteen studies based on a MOA for detecting communities in social networks. Each paper used the number of networks that were shown in Table 2 and commonly evaluated the performance of their algorithm using two metrics. This section demonstrates the detailed information of the common networks and explains metrics that are used to evaluate algorithms.

4.1. Common Networks

The networks that are used by each study consist of a different number of nodes and edges. Each node is strongly connected with others in the same community and weakly with other nodes in the different community. The connections between nodes are called edges [7]. This number of edges and nodes defines the size of the networks. Furthermore, in some of these networks, the number of communities was detected using a ground truth [33]. Table 3 shows the properties of some common networks that are used by the papers.

4.2. Metrics for Evaluating the Results

Q modularity and normalized mutual information (NMI) are two main metrics that are used to evaluate community detection in social networks. All of the studies in this review paper used these two metrics to evaluate the result of detecting communities for each network that was used in their paper.

TABLE 3: Some different size of common real-world networks

Networks	Number of nodes	Number of edges	Number of communities
Karate [34]	34	78	2
Dolphin [35]	62	159	2
Polbooks [36]	105	441	3
Football [37]	115	613	12
Citeseer [38]	3,327	4,676	6
Ca-GrQc [39]	5,242	14,496	-
CA-HelpTh [39]	9,877	25,998	-
Facebook [12]	4,039	88,234	-
Ca-HepTh2 [39]	12,008	118,521	-
Ca-AstroPh [39]	18,772	198,110	-
ca-CondMat [40]	23,133	93,497	-
Email-Enron [41]	36,692	183,831	-

The modularity Q, which can be computed without knowledge of the actual community labels of a network, was chosen as a measure of quality for the communities; a higher Q value indicates better performance in community detection [42]. The formula of modularity Q was defined as below.

$$Q = \sum_{i=1}^{k} [\frac{l_s}{M} - (\frac{d_s}{2M})^2]$$

k and M show the number of detected communities and edge's number in a network, respectively. Then, l_s is a number of edges for each node in the community i, and d_s is a total degree of nodes in the same community [42].

However, using truth grounds, the NMI was used in order to measure how similar the detected and real communities were. Greater value of NMI shows better performance in detecting communities [43]. The following is the definition of the NMI formula.

$$NMI(A,B) = \frac{\sum_{i=1}^{C^A}\sum_{j=1}^{C^B} C_{i,j} Log(\frac{C_{i,j}N}{C_{i.}C_{.j}})}{\sum_{i=1}^{C^A} C_{i.} \log(C_{i.}/n) + \sum_{j=1}^{C^B} C_{.j} \log(C_{.j}/n)}$$

In partitions A and B, CA and C^B present the number of communities, C shows the confusion matrix, and $C_{i,j}$ is a shared node between community i of A and community j of B. $C_{i.}$ or $C_{.j}$ determines the total elements of C in row i or column j, while the number of nodes in the network is n [43].

5. COMPARISON AND DISCUSSION

5.1. Compare Objectives used by Algorithms

All of the studies that were selected in this review paper used a MOA for detecting communities in the networks. Moreover, most of them used the MOA for optimizing two objectives instead of one that optimized three objectives. The number of objectives and their types are explained in Table 4. Furthermore, Fig. 3 explains the most frequent objectives used by the algorithms.

According to Fig. 3, modularity is the most frequently used by the algorithms. Then RC comes as the most commonly used after modularity. The modularity is used to measure the quality of network partitions and a RC is a solution to achieve as few as possible connections between the communities [21].

5.2. Problem Definition of Community Detection and Limitations

5.2.1. Overlapping communities

Refer to a problem in which one node simultaneously belongs to multiple communities [44]. Most of the studies in this review paper solved the mentioned problem. Fig. 4 shows the overlapping communities in the network.

TABLE 4: Demonstrates the number and type of objectives that optimized by the papers in the literature

Paper references	Number of objectives used by each paper		
	Objective 1	Objective 2	Objective 3
[18]	Optimize modularity	Optimize community score	
[19]	Maximizing modularity	Optimizing of frustration function	
[20]	Minimize KKM	Minimize RC	
[21]	Maximizing community score	Maximizing community fitness	Maximizing modularity
[22]	Minimizing the conductance of a community	Maximizing the number of key nodes in a community	
[23]	Optimize modularity	NMI	
[24]	Optimize NRA	Optimize RC	
[25]	Optimize community structure	Optimize community fitness	
[26]	Minimize KKM	Minimize RC	
[27]	Maximizing modularity	Minimizing the number of misclassified nodes	
[28]	Optimize KKM	Optimize RC	
[29]	Maximal clique detection	Community structure optimization	
[30]	Minimize community score	Minimize community fitness	
[31]	Modularity optimization	NMI	
[32]	Accuracy of community detection	Computational efficiency	

KKM: Kernel K-means, RC: Ratio cut, NMI: Normalized mutual information

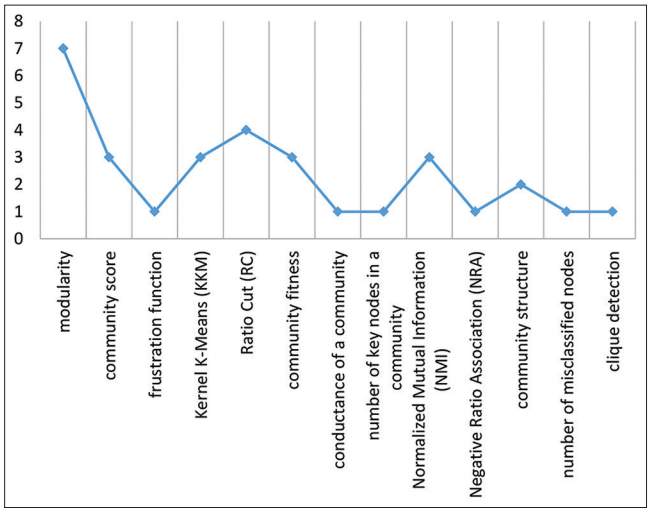


Fig. 3. The most frequent objectives used by the algorithms.

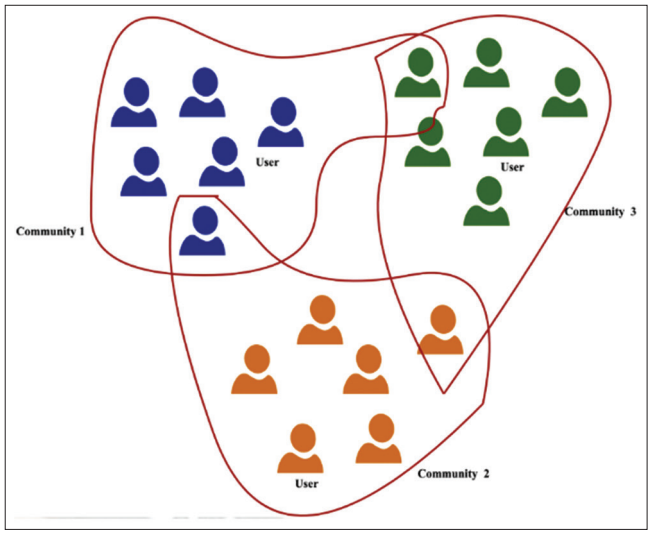


Fig. 4. Overlapping communities [45].

5.2.2. Dynamic networks

Is a network that changes structure over time. Either through the addition, removal, or change of nodes, edges, or both, dynamic networks differ from static networks with fixed nodes and edges [46]. Today, detecting communities in dynamic networks is the most popular challenge especially in terms of real-time community detection or recommender systems. Some algorithms try to solve this problem but not completely. It's a main gap in today's approaches. Any changes in the network the algorithm needs to execute again and it's a main limitation in the community detection algorithms.

5.2.3. Computational complexity

One of the most popular problems in MOAs for detecting communities is computational complexity because they must optimize two or three objectives simultaneously. This problem appears more while the network becomes larger over time. Furthermore, it's the main limitation in the majority of algorithms. Some algorithms simplified the complexity but the quality of the detection decreased and vice versa.

5.3. Addressing Limitations

5.3.1. Overlapping communities

Nowadays, the majority of algorithms prevent this problem by:

1. Evaluating the algorithm by robust functions Q modularity and NMI
2. Use high-accurate objectives, such as K-means [47], RC [48] and normalized cut [49]
3. Using evolutionary algorithms, such as NSGAII.

5.3.2. Dynamic networks problem

1. Where a new node is added, try to update only the nearby community instead of the entire network
2. Integrating graph neural network [50] to predict how communities in the network will change in the future.

5.3.3. Computational complexity

1. Instead of using random partitioning in the initialization process, you can use smart partitioning, such as Louvain [51] or Kernighan and Lin [52]
2. Develop a robust hybrid algorithm to reduce computational time
3. Improve mutation and crossover operators
4. Can be used early stopping to terminate the algorithm when no significant improvement is obtained over the generations
5. Used the compression technique over large-size networks.

5.4. Challenges and Future Work

According to the limitations that were explained in the previous sub-section, there are many challenges and future works.

1. Prepare a robust hybrid algorithm by using the specialty of the previous algorithms. That cover both the quality of detection and simplifies the complexity
2. Applied the algorithm to perform real-time community detection of the huge networks

3. Applied the algorithm for recommendation using the recommender system.

6. CONCLUSION

Nowadays, the MOA is the most widespread study for detecting communities in complex networks. The researchers have proposed various algorithms to achieve the highest accuracy and efficiency in detecting communities among different networks. Those approaches used MOAs to solve two or three conflict problems in the structure of the network. This research establishes a foundational understanding of various aspects of a MOA for community detection in complex networks by providing a detailed overview of the published literature. Our review paper focuses on novelties, development, processing, and published literature reviews, with the aim of identifying research gaps in the MOAs for detecting communities in complex networks. This survey summarized 15 different studies that were published in the years between 2019 and 2024. Furthermore, the methodology and contribution for each study have been demonstrated. After that, the datasets and evaluation metrics that were used by each approach have been presented. Then, discussion about the limitation and solution has been made.

The future work is developing a hybrid algorithm based on the gaps in the recent algorithms to detect more accurate community detection in complex networks.

REFERENCES

[1] S. E. De Leon-Aldaco, H. Calleja and J. Aguayo Alquicira. "Metaheuristic optimization methods applied to power converters: A review". *IEEE Transactions on Power Electronics*, vol. 30, no. 12, pp. 6791-6803, 2015.

[2] A. M. Nassef, M. A. Abdelkareem, H. M. Maghrabie and A. Baroutaji. "Review of metaheuristic optimization algorithms for power systems problems". *Sustainability*, vol. 15, no. 12, p. 9434, 2023.

[3] K. Diao, X. Sun, G. Bramerdorfer, Y. Cai, G. Lei and L. Chen. "Design optimization of switched reluctance machines for performance and reliability enhancements: A review". *Renewable and Sustainable Energy Reviews*, vol. 168, p. 112785, 2022.

[4] R. H. Stewart, T. S. Palmer and B. DuPont. "A survey of multi-objective optimization methods and their applications for nuclear scientists and engineers". *Progress in Nuclear Energy*, vol. 138, p. 103830, 2021.

[5] S. Chand and M. Wagner. "Evolutionary many-objective optimization: A quick-start guide". *Surveys in Operations Research and Management Science*, vol. 20, no. 2, pp. 35-42, 2015.

[6] K. Deng. "An efficient multi-objective community detection algorithm in complex networks". *Tehnicki vjesnik - Technical Gazette*, vol. 22, no. 2, pp. 319-328, 2015.

[7] P. Bedi and C. Sharma. "Community detection in social networks". *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, vol. 6, no. 3, pp. 115-135, 2016.

[8] B. S. Khan and M. A. Niazi. "Network community detection: A review and visual survey". arXiv preprint arXiv:1708.00977, 2017.

[9] A. Clauset, M. E. J. Newman and C. Moore. "Finding community structure in very large networks". *Physical Review E-Statistical, Nonlinear, and Soft Matter Physics*, vol. 70, no. 6, pp. 1550-2376, 2004.

[10] M. Rosvall and C. T. Bergstrom. "Maps of random walks on complex networks reveal community structure". *Proceedings of the National Academy of Sciences*, vol. 105, no. 4, pp. 1118-1123, 2008.

[11] A. L. Barabási. "Scale-free networks: A decade and beyond". *Science*, vol. 325, no. 5939, pp. 412-413, 2009.

[12] S. Fortunato. "Community detection in graphs". *Physics Reports*, vol. 486, no. 3-5, pp. 75-174, 2010.

[13] M. E. J. Newman. "Communities, modules and large-scale structure in networks". *Nature Physics*, vol. 8, no. 1, pp. 25-31, 2012.

[14] I. A. Doush, W. A. B. Alrashdan, M. A. Al-Betar and M. A. Awadallah. "Community detection in complex networks using multi-objective bat algorithm". *International Journal of Mathematical Modelling and Numerical Optimisation*, vol. 10, no. 2, pp. 123-140, 2020.

[15] Z. Beheshti and S. M. H. Shamsuddin. "A review of population-based meta-heuristic algorithms". *International Journal of Advances in Soft Computing and Its Applications*, vol. 5, no. 1, pp. 1-35, 2013.

[16] R. K. Congram, C. N. Potts and S. L. van de Velde. "An iterated dynasearch algorithm for the single-machine total weighted tardiness scheduling problem". *INFORMS Journal on Computing*, vol. 14, no. 1, pp. 52-67, 2002.

[17] F. Glover and C. McMillan. "The general employee scheduling problem. An integration of MS and AI". *Computers and Operations Research*, vol. 13, no. 5, pp. 563-573, 1986.

[18] A. H. Hosseinian and V. Baradaran. "A multi-objective multi-agent optimization algorithm for the community detection problem". *Journal of Information Systems and Telecommunication*, vol. 6, no. 1, pp. 169-179, 2019.

[19] M. M. Abdulrahman, A. D. Abood and B. A. Attea. "An Enhanced Multi-Objective Evolutionary Algorithm with Decomposition for Signed Community Detection Problem". In: *Presented at the 2020 2nd Annual International Conference on Information and Sciences* (*AICIS*). Fallujah, Iraq, 2020.

[20] Z. Liu, Y. Ma and X. Wang. "A compression-based multi-objective evolutionary algorithm for community detection in social networks". *IEEE Access*, vol. 8, pp. 62137-62150, 2020.

[21] T. Shaik, V. Ravi and K. Deb. "Evolutionary multi-objective optimization algorithm for community detection in complex social networks". *SN Computer Science*, vol. 2, pp. 1-25, 2021.

[22] Y. Su, K. Zhou, X. Zhang, R. Cheng and C. Zheng. "A parallel multi-objective evolutionary algorithm for community detection in large-scale complex networks". *Information Sciences*, vol. 576, pp. 374-392, 2021.

[23] M. E. A. Alkhalec Tharwat, M. F. M. Fudzee, S. Kasim, A. A. Ramli and M. K. Ali. "Multi-objective NSGA-II based community detection using dynamical evolution social network". *International Journal of Electrical and Computer Engineering*, vol. 11, no. 5, pp. 4502-4512, 2021.

[24] S. Pérez-Peló, J. Sánchez-Oro, A. Gonzalez-Pardo and A. Duarte. "A fast variable neighborhood search approach for multi-objective community detection". *Applied Soft Computing*, vol. 112, p. 107838, 2021.

[25] A. D. Abbood, A. A. Bara'a, A. A. Hasan and R. K. Everson. "*A Multi-Objective Evolutionary Algorithm with Neighbour Nodecentrality for Community Detection in Complex Networks*". Researchsquare, Durham, NC, 2022.

[26] S. Yang, Q. Li, W. Wei and Y. Zhang. "A multi-objective evolutionary algorithm based on mixed encoding for community detection". *Multimedia Tools and Applications*, vol. 82, no. 9, pp. 14107-14122, 2023.

[27] W. Zhu, H. Li and W. Wei. "A two-stage multi-objective evolutionary algorithm for community detection in complex networks". *Mathematics*, vol. 11, no. 12, p. 2702, 2023.

[28] L. Zhang, H. Yang, S. Yang and X. Zhang. "A macro-micro population-based co-evolutionary multi-objective algorithm for community detection in complex networks [research frontier]". *IEEE Computational Intelligence Magazine*, vol. 18, no. 3, pp. 69-86, 2023.

[29] C. Li. "Multi-objective optimization overlapping community detection algorithm based on subgraph structure". *Frontiers in Computing and Intelligent Systems*, vol. 3, no. 3, pp. 110-112, 2023.

[30] L. Yu, X. Guo, D. Zhou and J. Zhang. "A multi-objective pigeon-inspired optimization algorithm for community detection in complex networks". *Mathematics*, vol. 12, no. 10, p. 1486, 2024.

[31] L. Cai, J. Zhou and D. Wang. "Two-stage multi-objective evolutionary algorithm for overlapping community discovery". *PeerJ Computer Science*, vol. 10, p. e2185, 2024.

[32] J. Sheykhzadeh, B. Zarei and F. Soleimanian Gharehchopogh. "Community detection in social networks using a local approach based on node ranking". *IEEE Access*, vol. 12, pp. 92892-92905, 2024.

[33] J. Yang and J. Leskovec. "Defining and Evaluating Network Communities based on Ground-truth". In: *MDS '12: Proceedings of the ACM SIGKDD Workshop on Mining Data Semantics*, pp. 1-8, 2012.

[34] W. W. Zachary. "An information flow model for conflict and fission in small groups". *Journal of Anthropological Research*, vol. 33, no. 4, pp. 452-473, 1977.

[35] D. Lusseau. "The emergent properties of a dolphin social network". *Proceedings of the Royal Society of London. Series B: Biological Sciences*, vol. 270, no. Suppl_2, pp. S186-S188, 2003.

[36] M. E. J. Newman. "Modularity and community structure in networks". *Proceedings of the National Academy of Sciences*, vol. 103, no. 23, pp. 8577-8582, 2006.

[37] M. Girvan and M. E. J. Newman. "Community structure in social and biological networks". *Proceedings of the National Academy of Sciences*, vol. 99, no. 12, pp. 7821-7826, 2002.

[38] K. D. Bollacker, S. Lawrence and C. L. Giles. "CiteSeer: An Autonomous Web Agent for Automatic Retrieval and Identification of Interesting Publications". In: *Proceedings of 2nd International ACM Conference on Autonomous Agents*, ACM Press, pp. 116-123, 1998.

[39] J. Leskovec, J. Kleinberg and C. Faloutsos. "Graph evolution: Densification and shrinking diameters". *ACM transactions on Knowledge Discovery from Data*, vol. 1, no. 1, p. 2-es, 2007.

[40] J. Leskovec and A. Krevl. 2014. "*SNAP Datasets: Large Network Dataset Collection*". Available from: https://snap.stanford.edu/data [Last accessed on 2024 Jul 15].

[41] J. Leskovec, K. J. Lang, A. Dasgupta and M. W. Mahoney. "Community structure in large networks: Natural cluster sizes and the absence of large well-defined clusters". *Internet Mathematics*, vol. 6, no. 1, pp. 29-123, 2009.

[42] M. E. J. Newman. "Fast algorithm for detecting community structure in networks". *Physical Review E-Statistical, Nonlinear, and Soft Matter Physics*, vol. 69, no. 6, p. 066133, 2004.

[43] L. Danon, A. Díaz-Guilera, J. Duch and A. Arenas. "Comparing community structure identification". *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2005, no. 09, pp. P09008-P09008, 2005.

[44] J. C. Devi and E. Poovammal. "An analysis of overlapping community detection algorithms in social networks". *Procedia Computer Science*, vol. 89, pp. 349-358, 2016.

[45] P. B. Jadhav and V. B. Burra. "Deep learning in social networks for overlappering community detection". *International Journal on Recent and Innovation Trends in Computing and Communication*, vol. 10, no. 12, pp. 35-43, 2022.

[46] Z. Liu, Y. Dong, X. Zhao and B. Zhang. "A dynamic social network data publishing algorithm based on differential privacy". *Journal of Information Security*, vol. 8, no. 4, pp. 328-338, 2017.

[47] J. MacQueen. "Some Methods for Classification and Analysis of Multivariate Observations". In: *Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability*. University of California Press, Berkeley, 1967.

[48] C. M. Fiduccia and R. M. Mattheyses. "A Linear-time Heuristic for Improving Network Partitions". In: *Papers on Twenty-five years of Electronic Design Automation*, pp. 241-247, 1988.

[49] J. Shi and J. Malik. "Normalized cuts and image segmentation". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 8, pp. 888-905, 2000.

[50] F. Scarselli, M. Gori, A. C. Tsoi, M. Hagenbuchner and G. Monfardini. "The graph neural network model". *IEEE Transactions on Neural Networks and Learning Systems*, vol. 20, no. 1, pp. 61-80, 2009.

[51] V. D. Blondel, J. L. Guillaume, R. Lambiotte and E. Lefebvre. "Fast unfolding of communities in large networks". *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2008, no. 10, p. P10008, 2008.

[52] B. W. Kernighan and S. Lin. "An efficient heuristic procedure for partitioning graphs". *The Bell System Technical Journal*, vol. 49, no. 2, pp. 291-307, 1970.

ORIGINAL RESEARCH ARTICLE UHD JOURNAL OF SCIENCE AND TECHNOLOGY

Small Dam Design and Construction for Sustainable Water Resources Management: A Comprehensive Review

Abdalmajeed Mohammed Rahman, Nawbahar Faraj Mustafa*

Department of Water Resources Engineering, University of Sulaimani, Sulaymaniyah, Iraq

ABSTRACT

Small dams are crucial in water resource management, particularly in regions with water scarcity and climate unpredictability. Despite their cost-effectiveness, the construction of small dams often lacks engineering standards, which raises concerns about their long-term stability and safety. This study reviews the design, construction, stability, and protection of small dams, emphasizing the importance of proper site selection, geological and hydrological studies, and advanced methodologies, such as Geographic Information Systems and multi-criteria decision-making approaches in dam evaluation. Furthermore, the study highlights the significance of detailed planning, material selection, and quality construction to ensure dam longevity. It also discusses the role of modern tools, such as HEC-HMS, HEC-RAS, and GeoStudio in assessing flood risks, seepage, and stability. Inadequate design, particularly in the face of extreme weather events, can lead to dam failures, emphasizing the need for comprehensive planning and rigorous assessments. Through an analysis of various studies and case examples, this paper aims to provide insights into sustainable small dam construction and water resources management practices that ensure their effectiveness and resilience in addressing water scarcity challenges.

Index Terms: Small Dam Construction, Water Resource Management, Stability Analysis, Construction Practices, Site Selection, Hydrology

1. INTRODUCTION

Small dams and reservoirs have consistently drawn the interest of rural development authorities and aided human development by providing reliable sources of water [1], [2]. Small dams are increasingly recognized for their role in water resource management, particularly in regions facing water scarcity and unpredictable climate patterns [3].

Access this article online

DOI: 10.21928/uhdjst.v9n1y2025.pp55-64

E-ISSN: 2521-4217

P-ISSN: 2521-4209

A "small" dam is typically defined as one where the maximum height above the original streambed does not exceed approximately 15 m (about 50 feet). In addition, it is characterized by a volume not large enough to justify the use of detailed design methods typically reserved for larger dams. Furthermore, if a low dam has a volume greater than 765,000 cubic meters, it would no longer be classified as small [4] as shown in Fig. 1, the Al Abila Dam is a small dam in Iraq's Western Desert [5].

A study of three small dams, Jawa, Kasala, and Dhok Sanday Mar, evaluated the impact of the small dams on agriculture and groundwater development in the state of Punjab in Pakistan and the analysis of inflow–outflow of the dams (Figs. 2-4) shows that if properly managed, the storage is sufficient to irrigate all the croplands within the command area [6].

Corresponding author's e-mail: Nawbahar Faraj Mustafa, Department of Water Resources Engineering, University of Sulaimani, Kirkuk Road, Sulaymaniyah, 46002, Iraq. E-mail: nawbahar.mustafa@univsul.edu.iq

Received: 05-02-2025 Accepted: 27-02-2025 Published: 27-03-2025

Fig. 1. Al Abila Dam [5].

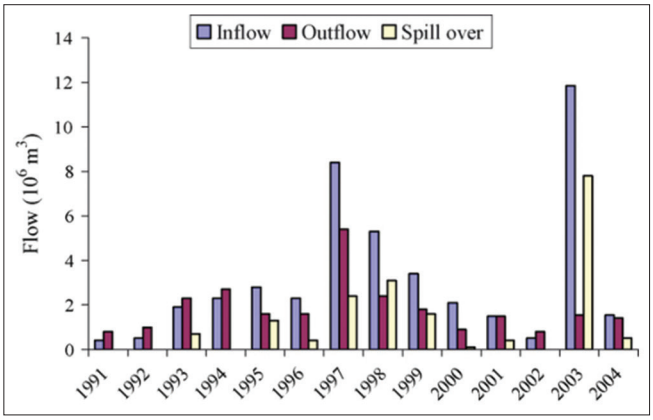


Fig. 2. Yearly inflow-out flow of Khasala dam.

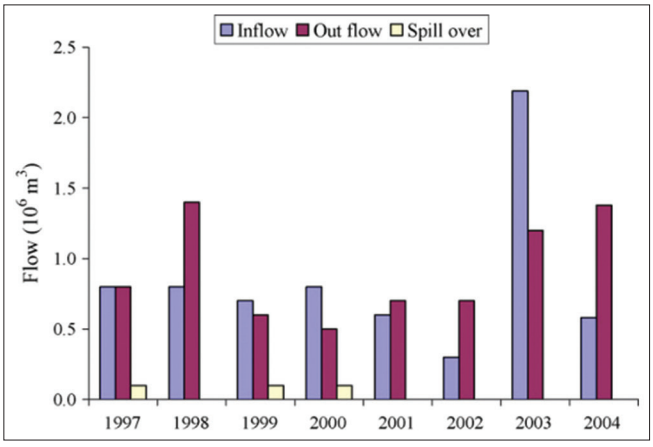


Fig. 3. Yearly inflow-outflow of Jawa dam.

Salimi *et al.* highlighted the dam's crucial role in irrigation, drinking water, and industrial use in the dry, semi-arid region. The need for coordinated, long-term water management for agricultural, domestic, and industrial purposes is emphasized [7]. The construction and maintenance of small dams often pose significant challenges [8], especially when built without engineering standards. Farmers frequently construct small dams without engineering input, which

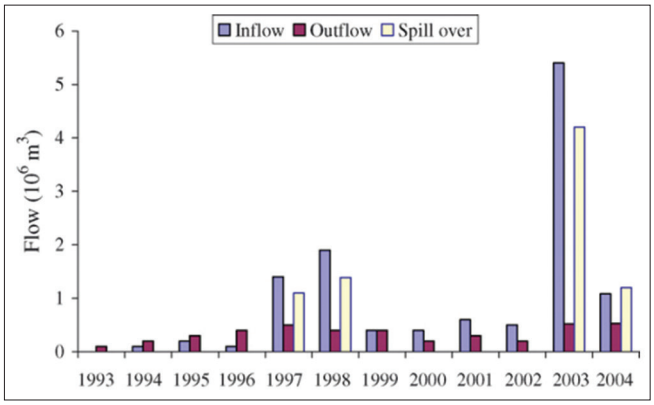


Fig. 4. Yearly inflow-outflow of D.S. Mar dam.

raises concerns about their long-term stability, safety, and effectiveness. Valuable insights can be gained from both the successes and failures of dam applications and operations. These can guide water resource managers, policymakers, and stakeholders in achieving sustainable development goals, particularly in climate change and growing water crises [3].

Dams can be grouped differently, especially considering their height and storage; the classification can be arranged as follows:

1. Low Dams: These are usually modest in height, standing <15 m (or about 50 feet) tall. They typically don't store much water, <765,000 cubic meters, or 1 million cubic yards, making them ideal for local purposes, such as supporting irrigation or providing drinking water to nearby communities.
2. Medium Dams: Medium dams are met when a bit of height is stepped up. These range from 15 to 30 m (50–100 feet) high, have a moderate storage capacity, and are versatile. They often supply regional water needs, support agricultural irrigation, or generate hydroelectric power.
3. High Dams: High dams tower over us at more than 30 m (or around 100 feet). These giants can store large volumes of water, often reaching millions of cubic meters. They play crucial roles in significant water management projects, including large hydroelectric power stations, extensive irrigation systems, and considerable flood control.

It's worth noting that while low and medium dams focus on local or regional needs, the high dams have a much broader impact. Each type plays a vital role in water management and sustainability, tailored to specific needs based on height and storage capacity [9].

The availability of dam fill materials influenced the dam design considerably [10]. Dams can also be classified based on the construction materials used. Embankment (earthen) Dams are dams constructed with material, such as compacted soil, clay, or rockfill [11]. Rockfill dams are made of rock fragments or gravel with an impervious core from clay or concrete [12] while concrete dams are made from concrete [13].

Despite challenges, small dams are considered cost-effective solutions for water management in rural areas, especially when conventional infrastructure is not feasible. As the world faces increasing pressures from climate change, the ability of these small dams to withstand extreme weather events, such as prolonged droughts, becomes more critical [14].

Aureli *et al.* [15] reviewed several factors contributing to dam failures, including overtopping, which often results from inadequate design considerations for flood capacity, and internal erosion due to piping. They also highlighted the importance of proper management practices to mitigate risks during the design and construction phases. These factors underline the critical need for robust pre-construction assessments and planning to prevent potential failures.

This study aims to review key primary research on small dam design, focusing on aspects, such as construction, stability, safety, performance, and sustainable water management. This area has not been extensively explored. The primary studies on dam design include site selection, geological and hydrological assessments, and environmental impact evaluations. The design methodology is examined through a comprehensive approach that addresses planning, material selection, construction quality, dam height, reservoir capacity, site condition assessment, structural components, hydraulic management, earthquake risk mitigation, spillway and bottom outlet design, physical modeling, numerical simulations, sediment estimation, dam sustainability, cost, operation, and management, as well as environmental impact and community protection.

2. DAM SITE SELECTION AND PRELIMINARY INVESTIGATION

Proper site selection is the first step in ensuring the stability and effectiveness of small dams. Traditional site selection methods rely on local knowledge and experience [16]–[18]. Still, modern approaches incorporate scientific methods and tools, such as Geographic Information Systems (GIS), remote sensing (RS), and hydrological modeling, and GIS-based Fuzzy Analytic Hierarchy Process (FAHP) and Adaptive Neuro-Fuzzy Inference System (ANFIS) approaches were applied to evaluate potential dam sites, integrate environmental, hydrological, and economic factors to identify the most suitable location for mitigating water scarcity issue [19]–[24]. Appropriate site selection factors include hydrological aspects, such as runoff potential, slope, soil type, and vegetation, and socioeconomic considerations, such as proximity to communities and agricultural land [21], [25]. The design phase should also incorporate an understanding of the local geology and soil conditions [26]. The investigation process for dam construction also involves verifying the quality and quantity of fill materials to confirm their properties [10]. The decision-making methods for optimizing dam site selection were also used, emphasizing a structured approach to evaluating potential reservoir locations based on various measurable characteristics [27], [28]. Hashim and Sayl explored methodologies for selecting dam construction and rainwater harvesting. The study utilized GIS techniques, employing the weighted linear combination (WLC) and Boolean overlay methods. By assigning scores and weights, WLC analyzed parameters, such as runoff depth, slope, soil texture, land use and land cover (LULC), and proximity to irrigated lands, roads, and residential areas. The Boolean overlay method used criteria, such as stream order and fault proximity for binary classification, providing a comprehensive approach to site suitability assessment [29]. In a study by [30], they emphasize the importance of proper site selection for dams and water harvesting to combat water scarcity and enhance agricultural productivity. Critical factors include slope, soil type, vegetation, rainfall, and runoff capacity. Drone-assisted topographic analysis can optimize site selection, ensuring sustainable water management and agricultural resilience.

Abualhaija and Mohammad analyzed the Kufranja Dam in Jordan to address water shortages in an arid climate with limited freshwater resources. They concluded that the key factors for successful dam construction include geological surveys, hydrological features, and environmental impacts on agriculture and urban development [31].

Wang *et al.* highlight the importance of scientific dam location and construction for irrigation, flood control, and hydroelectric power. They discuss GIS/RS, Multi-Criteria Decision Making (MCDM), and machine learning as vital siting methods, emphasizing hydrology, geology, and socio-economic factors. They resulted in neglecting these considerations will risk environmental damage and dam

failures [23]. Rather *et al.* examine the challenges of dam site selection in the Jhelum Basin, emphasizing that inadequate consideration of factors, such as landslide-prone areas and unstable soils can lead to dam failures. They stress the need to thoroughly evaluate topography, hydrology, geology, and socio-environmental factors. The study highlights that poorly located dams fail to manage floods effectively. The authors recommend adopting a multi-criteria analysis approach to address these issues to enhance site selection and prevent future failures [32].

In conclusion, various studies highlight the importance of selecting the proper site for small dams to ensure stability, effectiveness, and sustainable water management. Modern approaches integrate GIS, RS, hydrological modeling, and advanced decision-making methods, such as FAHP, ANFIS, WLC, and Boolean overlay. These methods assess runoff potential, slope, soil type, vegetation, proximity to communities and agricultural land, and the availability of quality fill materials. In addition, topographic analysis using drones has been proposed to optimize site selection. These researches underscore the necessity of addressing hydrological, geological, and socio-economic factors to mitigate challenges, such as water scarcity, population growth, climate change, and unstable soils. Case studies reveal that poorly selected sites risk dam failure, environmental damage, and ineffective flood management, emphasizing the need for comprehensive, multi-criteria analysis to enhance decision-making and improve water infrastructure resilience.

Numerous studies have been conducted in Iraq on dam site selection using different parameters [33]–[36]. For example, Al-Ansari *et al.* tested two areas (northwestern and northeastern parts of Iraq) for the feasibility of WH using small dams not more than 6 m in height. They identified suitable sites for small dam construction based on different parameters and applied the Watershed Modeling System and linear programming optimization techniques [37]. Another study to select optimal dam sites was conducted in the Deewana watershed (North of Iraq) using a combination of RS and GIS with multi-criteria decision analysis models. As a result, they identified (5.55%) and (21.81%) as highly suitable for constructing dams in the dam site selection maps and found that 11 proposed dam sites are ideal for dam construction [38].

3. GEOLOGY STUDY

Studying geology in dam design is essential for ensuring the structure's stability and safety by analyzing site conditions, foundation properties, and construction materials [39]. It involves geological mapping, geophysical surveys, and borehole sampling to evaluate rock and soil quality, permeability, and bearing capacity. Geological hazards, such as earthquakes, landslides, and karst features are assessed to mitigate risks. In addition, the availability and durability of local materials are tested for suitability in construction [40]. Tools, such as GeoStudio and GIS aid in seepage analysis, slope stability, and site mapping, ensuring a safe and efficient dam design tailored to the geological context [41].

Sissakian *et al.* highlighted the importance of geological investigations in dam construction, using the Mosul Dam to show the risks of inadequate assessments. Before site selection, they emphasized evaluating foundation geology, reservoir integrity, slope stability, and material availability. The process includes three steps: preliminary investigations for general insights, initial design studies for detailed data, and final design assessments for engineering needs. Key recommendations include geological mapping, drilling, and hydrogeological studies. Misinterpreting geological data, such as karstification and gypsum behavior, led to challenges at Mosul Dam, underscoring the need for thorough, expert-led evaluations to ensure dam safety [42].

Birhanu *et al.* evaluated the Upper Guder Dam slope stability in Central Ethiopia, addressing critical issues related to small dam construction, such as site preparation, material selection, spillway design, and costs. The results show that geological formations are susceptible to leakage and instability due to complex discontinuities and geological variances, particularly in the left abutment, which poses wedge failure risks. In contrast, increased pore water pressure compromises the right abutment's stability under saturated conditions. The study emphasized the importance of comprehensive geological assessments during the design and construction phases to prevent failures [43].

In summary, Geological studies are essential in dam design to ensure stability and safety by analyzing site conditions, foundation properties, and material suitability. These involve mapping, surveys, and sampling to assess hazards, such as landslides, earthquakes, and karst features. Tools, such as GeoStudio and GIS aid in seepage analysis and slope stability. Misinterpreted geological features have caused failures in past projects, highlighting the need for thorough, expert-led assessments. Comprehensive evaluations during design and construction address issues, such as material selection, spillway design, and abutment stability, ensuring safe and efficient dam construction.

4. HYDROLOGICAL STUDY

The hydrological study is essential for estimating peak floods with different return periods and average annual runoff that reaches the dam site to fill the reservoir. Moreover, assessing the sediment volume is considered in the hydrologic study [44]–[46]. Hydrological modeling tools, such as HEC-HMS and HEC-RAS, are increasingly used to predict peak discharges, assess flood risks, and evaluate dams' hydraulic performance. These tools help engineers determine the optimal size and structure of a dam to accommodate anticipated flood events while minimizing the risk of failure [47], [48].

Othman *et al.* applied a MCDM approach to assess dam sites, incorporating factors, such as soil type, elevation, precipitation, and proximity to faults. They highlighted the need for thorough geological and hydrological surveys, as past infrastructure suffered from inadequate evaluations [49].

Thieme *et al.* [50] highlight the importance of considering ecological impacts and hydrological changes during dam planning and construction. They note a rise in protected areas being downgraded or removed due to dam construction. They recommend avoiding dam construction near protected areas and removing problematic dams where possible. The paper calls for better alignment between development goals and ecological conservation to ensure functional freshwater ecosystems and urges policymakers to manage freshwater ecosystems responsibly and promote measures to restore rivers.

Hydrologic studies estimate peak floods, runoff, and sediment at dam sites, ensuring safe and efficient designs. Tools, such as HEC-HMS and HEC-RAS optimize dam performance and flood risk management. Surveys and multi-criteria approaches highlight the importance of hydrological considerations, urging avoidance of protected areas and addressing common dam failures, such as overtopping and erosion. Cascade systems require precise planning to manage peak flows and ensure effective flood mitigation.

5. DESIGN METHODOLOGIES

Small dams are vital in managing water resources, particularly in areas with irregular or limited rainfall. Historically, they have been indispensable for agriculture and ensuring food security. However, failures often occur due to poor design and construction deficiencies [51]. Once a dam site is preliminarily selected, conducting a detailed survey and developing area-volume-elevation curves are essential [18]. These are followed by adopting design methodologies tailored to site-specific conditions to ensure dam safety [52]. These methodologies involve analyzing geological, hydrological, and structural factors to optimize the dam type, foundation stability, and seepage control systems [43], [53].

Říha *et al.* [54] discussed small dam construction and operational dynamics, focusing on the cascade system along the Cižina River in the Czech Republic. Various construction methods were described, emphasizing the homogeneous nature of the dams built with clay and detailing their dimensions, materials, and emergency spillway configurations, which are critical for managing flood waters and preventing overtopping failures. Stability analysis is conducted by simulating breaching scenarios, considering overtopping and piping as potential failure modes. Seepage analysis was also critical, highlighting the vulnerability of small dams to erosion due to internal seepage pressures. The article cited that the most common causes of failure stem from overtopping and internal erosion, which is further complicated by insufficient spillway capacity compared to incoming floods.

The key aspects, according to numerous studies on dam design, can be included as follows:

1. Comprehensive Planning and Modeling: Thorough planning and advanced modeling are essential to address site-specific challenges, predict risks, and ensure dam safety and functionality [2], [55].
2. Material Selection and Construction Quality: Choosing appropriate materials, proper layering, and maintaining high construction standards are critical for ensuring the dam's stability and longevity [55]–[61]. Proper layering of materials, such as shell materials, gabions, cores, and filters is essential for ensuring stability and resisting seepage and erosion [55], [62], [63].
3. Dam Height and Reservoir Capacity: Determining the dam's height and reservoir capacity is crucial for meeting water storage needs while maintaining structural safety [56].
4. Foundation Evaluation: Assessing foundation stability [58], [64], [65], identifying potential seepage paths is essential for ensuring the dam's structural integrity and long-term performance [56], [66].
5. Structural Components and Hydraulic Management: Upstream and downstream slopes, spillways, and bottom outlets are vital in managing hydraulic pressures and preventing dam failures [57], [63], [64], [67].
6. Resilience Against Natural Hazards: Mitigating earthquake risks [2], [68] and extreme weather involves

robust spillway systems to manage water flow and prevent overtopping. Advanced tools, such as HEC-RAS and RS aid rainfall-runoff analysis, hydraulic modeling, and peak discharge predictions, enhancing resilience and minimizing community impact [55], [56].

7. Spillway and Bottom Outlet Design: The design of spillways and bottom outlets is crucial for managing water flow, preventing overtopping, and ensuring the safe release of excess water during high-flow events [55], [58], [60], [61], [63]. Computational fluid dynamics (CFD) techniques refine spillway design and performance evaluation [45]. While 3D finite element models analyze internal mechanics, including soil layering, material properties, and hydraulic interactions [69].
8. Physical Models and Numerical Simulations: Physical models and numerical simulations [70] combine experimental and numerical analyses to evaluate seepage characteristics and structural displacements under various loading conditions [62], [65]. Stability and seepage analysis is crucial for preventing failure risks associated with overtopping and piping, significant causes of dam failure [68]. Advanced tools, such as GeoStudio aid in seepage analysis, slope stability, and stress evaluation, incorporating factors, such as water pressure, seismic activity, and settlement risks [71]. The dam's composition, geometric design, and surrounding vegetation enhance stability by limiting erosion and improving resilience during peak precipitation events [69]. SEEP/W software analyzes hydraulic performance aspects, including clay core thickness, cutoff wall depth, uplift pressure, hydraulic gradient, and seepage discharge [72]. The clay cores and cutoff walls are essential in managing seepage, noting that their placement significantly affects hydraulic gradients and discharge rates [58]–[60], [73]. Finite element methods are used to analyze the dam's stability and seepage characteristics [74].
9. Sediment Estimation and Dam Sustainability: Sediment estimation in small dams is essential for assessing long-term functionality. Sedimentation reduces storage capacity, affecting water retention and flood control efficiency [75], [76]. Embaye *et al.* stressed the significance of stability analysis for small dams, focusing on slope assessments, spillway design, and bottom outlets to manage erosion, water flow, and sediment buildup. Challenges, such as sediment accumulation and weak water user associations were identified as key contributors to dam failures [77]. Abdullah *et al.* examined the status and opportunities for water harvesting in Iraq. They noted that sediment accumulation is a major challenge to dam functionality, especially in the Eastern Valleys, where high sediment loads reduce reservoir capacities. Strategies are employed to minimize soil erosion and address potential risks of failure caused by sediment accumulation or structural deficiencies [4].
10. Cost, Operation, and Management: The cost of dam construction and operation, along with regular monitoring and effective management, are crucial for ensuring long-term functionality and safety. Continuous assessment and maintenance help detect potential issues early, preventing costly failures and optimizing performance [2], [59], [66], [78].
11. Material Selection and Construction Techniques: The analysis emphasizes the importance of carefully selecting materials and construction techniques to prevent failure, including controlling moisture content in the core material and ensuring proper drainage to manage seepage. Appropriate design and execution of these elements are vital for the dam's long-term stability and performance in irrigation management [9]. The importance of drainage methods, such as toe or horizontal drains, to reduce seepage and improve dam resilience is emphasized in the study, with support provided for earth-fill dam design to ensure sustainability and safety [70]. Controlling the permeability ratio between the shell and core layers and using coarse-grained filters to manage seepage and prevent erosion are recommended [74].
12. Environmental Impact and Community Protection: Protecting communities and minimizing environmental impact are key considerations in dam design and management [2], [66], [74]. Kondolf and Yi [78] focus on dam renovation strategies to extend reservoirs' operational lifespan and mitigate adverse effects on river ecosystems. They discuss minor dam construction aspects, including structural retrofits, fish passage devices, modifications for improved environmental flow, and sustainable sediment management practices. The study advocates enhanced design and operation to meet ecological standards, concluding that proactive renovation efforts are crucial for enhancing reservoir sustainability and environmental functionality.

As a result, the design and management of small dams require comprehensive planning, material selection, and construction quality to ensure stability and longevity. Key aspects include site-specific evaluations, proper layering of materials, such as clay cores, and addressing seepage and erosion risks. Structural components, such as spillways, bottom outlets, and dam slopes are essential for managing hydraulic pressures and preventing failure. Advanced tools, such as HEC-RAS, CFD,

and GeoStudio are used for modeling, stability analysis, and seepage management. Sediment accumulation is a significant challenge, reducing storage capacity and efficiency, especially in areas with high sediment loads. Regular monitoring and effective management help prevent costly failures, while sustainable practices, such as retrofitting and sediment flushing, can mitigate environmental impacts.

6. CLIMATE CHANGE AND RISK MANAGEMENT

Climate and socioeconomic changes continue to pose challenges to developing effective adaptation measures [79]. Climate change adds complexity to the design and management of small dams. Increased rainfall frequency and intensity and shifting temperature patterns can alter hydrological conditions in dam catchments, impacting runoff patterns and raising flood risks. Dam failure risk rises due to climate change, with more frequent and intense rainfall causing floods that exceed design capacities. Climate models predict that the return period of extreme flooding events will be reduced [80]. To mitigate these risks, it is recommended that dam design and operational protocols be updated, flood criteria be re-evaluated, regular inspections be conducted using updated hydrological models [81], and additional flood control infrastructure be constructed. Incorporating climate change into risk assessments ensures dams' long-term safety [82], protecting downstream communities, and enhancing water resource management [83].

7. RESULTS AND DISCUSSION

It is crucial to select proper dam sites for sustainable water resources management. New applied techniques include using GIS, RS, and hydrological modeling and integrating FAHP, ANFIS, and WLC. The main key factors confirmed by the studies are the runoff potential, rainfall depth, peak flood, topography slope, elevation, soil type and fault, LULC, and socio-economic aspects, such as proximity to communities, roads, and infrastructure. In contrast, neglecting these can lead to dam failure and environmental risk.

Focus is put on the geological studies that are crucial for dam design, ensuring stability through assessing site conditions, foundation, and material suitability. This includes geological mapping, evaluating rock and soil samples, testing foundation permeability and bearing capacity, and analyzing earthquake hazards. Seepage and slope stability analysis using GeoStudio are also essential. Case studies highlight failures due to misunderstanding geology in ensuring efficient dam construction. Effective reservoir management requires hydrological study assessments to estimate the peak floods, average annual runoff, and sediment volume at dam outlets; tools, such as HEC-HMS and HEC-RAS assist in assessing these parameters. MCDM approaches help to integrate hydrological and geological site selection factors. In addition, ecological considerations highlight the need not to use protected areas. Proper hydrological assessments and environmental impact mitigation ensure efficient dam performance.

The effectiveness of small dam design depends on good surveys and accurate design methodologies. Key design considerations include foundation stability, material selection, seepage control, and hydraulic management. Common overtopping failures highlight the need for robust spillway designs. Sediment estimation is critical for long-term functionality, while cost-effective construction and operational management ensure sustainability. Combining physical models and numerical simulations enhances stability assessments, improving small dams' resilience and efficiency. Climate change poses additional risks, altering hydrological patterns and increasing flood frequency, which should be considered in dam design and management.

8. CONCLUSIONS

Small dams are increasingly recognized as crucial for water resource management, especially in areas facing water scarcity and unpredictable climate conditions. While these dams offer a cost-effective solution, they present challenges, mainly when constructed without proper engineering practices. Such dams' stability, safety, and long-term functionality are key concerns, mostly when built by farmers without professional input. Despite these challenges, small dams remain essential in rural areas, providing vital services, such as irrigation, groundwater recharge, and flood control.

Research into small dam construction and performance highlights the importance of proper site selection, geological studies, and hydrological assessments to ensure these structures' safety and stability. Modern methods, including GIS, RS, and hydrological modeling, significantly improve site selection and dam design, aiding in more informed decision-making and reducing risks from poor construction practices. MCDM approaches, which assess hydrological, geological, and socio-economic factors, ensure these dams are resilient to climate change, water scarcity, and population pressures.

Geological studies are crucial for understanding site conditions, material suitability, and potential hazards, such as landslides or

earthquakes. Tools, such as GeoStudio and GIS help engineers assess the seepage, slope stability, and material properties, thereby minimizing the risk of dam failure. Hydrological studies, including flood peak runoff and sediment volume estimations, guide the design of dams that can withstand extreme weather events and operate safely over time.

Comprehensive planning, material selection, structural analysis, and resilience strategies are key to effective dam design. Advanced tools, such as HEC-HMS, HEC-RAS, and CFD improve dam safety by addressing overtopping, internal erosion, and seismic risks. Proper design of spillways, bottom outlets, and other structural components is critical for preventing failures, while numerical simulations and physical models offer insights into dam behavior under varying conditions.

Adopting modern site selection techniques, conducting thorough geological and hydrological assessments, adhering to engineering standards, and ensuring regular maintenance and inspections are essential for improving the safety and longevity of small dams. Ultimately, small dams' success hinges on careful planning, proper construction, and continuous monitoring, while incorporating climate change risks into design and management practices ensures their resilience in the face of evolving challenges.

REFERENCES

[1] J. Payen, J. M. Faurès, and D. Vallée, "Small reservoirs and water storage for smallholder farming--the case for a new approach," *Gates Open Research,* vol. 3, no. 387, p. 387, 2019.

[2] E. Umukiza, K. F. Abagale, and T. A. Adongo, "A Review on A review on significance and failure causes of small-scale irrigation dams in arid and semi-arid lands," *Journal of Infrastructure Planning and Engineering (JIPE),* vol. 2, no. 2, pp. 1-9, 2023.

[3] A. N. Angelakis, A. Baba, M. Valipour, J. Dietrich, E. Fallah-Mehdipour, J. Krasilnikoff, E. Bilgic, C. Passchier, V. A. Tzanakakis, R. Kumar, Z. Min, N. Dercas, and A. T. Ahmed, "Water Dams: From ancient to present times and into the future," *Water,* vol. 16, no. 13, p. 1889, 2024.

[4] M. Abdullah, N. Al-Ansari, and J. Laue, "Water harvesting in Iraq: Status and opportunities," *Journal of Earth Sciences and Geotechnical Engineering,* vol. 10, no. 1, pp. 199-217, 2020.

[5] A. Adham, S. Seeyan, R. Abed, K. Mahdi, M. Riksen, and C. Ritsema, "Sustainability of the Al-Abila Dam in the Western Desert of Iraq," *Water,* vol. 14, no. 4, p. 586, 2022.

[6] M. Ashraf, M. A. Kahlown, and A. Ashfaq, "Impact of small dams on agriculture and groundwater development: A case study from Pakistan," *Agricultural Water Management,* vol. 92, no. 1-2, pp. 90-98, 2007.

[7] N. Salimi, A. Feizi, S. Rasinezami, and A. Kanooni, "Management and planning of water resources allocation at the scenario analysis using system dynamics model: A case study on Yamchi dam basin, Iran," *Journal of Applied Research in Water and Wastewater,* vol. 8, no. 1, pp. 14-20, 2021.

[8] M. B. Gutti, J. A. Yunusa, U. Busuma, and S. A. Wazir, "Navigating the depths: An assessment of the management and operations of Alau Dam reservoir," *International Journal of Civil and Structural Research,* vol. 12, no. 2, pp. 1-14, 2024.

[9] C. P. Antonopoulos and T. C. Triantafillou, "Analysis of FRP-strengthened RC beam-column joints," *Journal of Composites for Construction,* vol. 6, no. 1, pp. 41-51, 2002.

[10] M. Ozcelik, "Effects of construction material properties on dam type selection (Sariguzel Dam–Turkey)," *International Journal of Geotechnical Engineering,* vol. 11, no. 4, pp. 368-375, 2017.

[11] A. Penman, "On the embankment dam," *Geotechnique,* vol. 36, no. 3, pp. 303-348, 1986.

[12] V. Andjelkovic, N. Pavlovic, Z. Lazarevic, and S. Radovanovic, "Modelling of shear strength of rockfills used for the construction of rockfill dams," *Soils and Foundations,* vol. 58, no. 4, pp. 881-893, 2018.

[13] M. E. Emiroglu, "Influences on selection of the type of dam," *International Journal of Science and Technology,* vol. 3, no. 2, pp. 173-189, 2008.

[14] I. Chirisa, F. Madya, R. Katsande-Ncube, N. Ndemo, and G. Mhlanga, "Dam Construction and the Establishment of New Ecosystems in Arid Places," in *The Palgrave Encyclopedia of Sustainable Resources and Ecosystem Resilience*. Springer, Germany, pp. 1-14, 2024.

[15] F. Aureli, A. Maranzoni, and G. Petaccia, "Review of HistoricalDam-Break Events and Laboratory Tests on Real Topographyfor the Validation of Numerical Models," *Wate*r, vol. 13, no. 14, 2021.

[16] B. A. Yifru, M. G. Kim, J. W. Lee, I. H. Kim, S. W. Chang, and I. M. Chung, "Water storage in dry riverbeds of arid and semi-arid regions: Overview, challenges, and prospects of sand dam technology," *Sustainability,* vol. 13, no. 11, p. 5905, 2021.

[17] A. S. Ibrahim, I. S. Al Zayed, F. S. Abdelhaleem, M. M. Afify, A. Ahmed, and I. Abd-Elaty, "Identifying cost-effective locations of storage dams for rainfall harvesting and flash flood mitigation in arid and semi-arid regions," *Journal of Hydrology: Regional Studies,* vol. 50, p. 101526, 2023.

[18] T. Stephens, *Manual on Small Earth Dams: A Guide to Siting, Design and Construction,* FAO, Rome, 2010.

[19] E. Abushandi and S. Alatawi, "Dam site selection using remote sensing techniques and geographical information system to control flood events in Tabuk City," *Hydrology Current Research,* vol. 6, no. 1-1000189, pp. 1-13, 2015.

[20] R. Al-Ruzouq, A. Shanableh, A. G. Yilmaz, A. Idris, S. Mukherjee, M. A. Khalil, and M. B. A. Gibril, "Dam site suitability mapping and analysis using an integrated GIS and machine learning approach," *Water,* vol. 11, no. 9, p. 1880, 2019.

[21] N. F. Mustafa, S. F. Aziz, H. M. Ibrahim, K. Z. Abdulrahman, J. T. Abdalla, \and Y. A. Ahmad, "Double Assessment of Dam Sites for Sustainable Hydrological Management Using GIS-Fuzzy Logic and ANFIS: Halabja Water Supply Project Case Study," *Iranian Journal of Science and Technology, Transactions of Civil Engineering*, pp. 1-19, 2024.

[22] A. M. Noori, B. Pradhan, and Q. M. Ajaj, "Dam site suitability assessment at the Greater Zab River in northern Iraq using remote sensing data and GIS," *Journal of Hydrology,* vol. 574, pp. 964-979, 2019.

[23] Y. Wang, Y. Tian, and Y. Cao, "Dam siting: A review," *Water,* vol. 13, no. 15, p. 2080, 2021.

[24] K. N. Gharib, N. F. Mustafa, and H. M. Rashid, "Urban rainwater harvesting assessment in Sulaimani heights district, Sulaimani City, KRG, Iraq," *UHD Journal of Science and Technology,* vol. 5, no. 1, pp. 48-55, 2021.

[25] H. Darabi, E. Moradi, A. A. Davudirad, M. Ehteram, A. Cerda, and A. T. Haghighi, "Efficient rainwater harvesting planning using socio-environmental variables and data-driven geospatial techniques," *Journal of Cleaner Production,* vol. 311, p. 127706, 2021.

[26] F. Katty, B. Ankidawa, G. Obiefuna, Y. Valdon, and I. Kwami, "A proposed dam site investigation using integrated geophysical and geotechnical methods; A case study of Loko area, Northeastern Nigeria," *Global Journal of Geological Sciences,* vol. 22, no. 1, pp. 31-45, 2024.

[27] I. Gacko, Z. Muchová, Ľ. Jurík, K. Šinka, L. Fabian, and F. Petrovič, "Decision making methods to optimize new dam site selections on the Nitra River," *Water,* vol. 12, no. 7, p. 2042, 2020.

[28] Y. N. Kontos, I. Bampekos, and K. L. Katsifarakis, "Multi-criteria optimization approach to the design of small dams' systems along mountainous stream beds," *Journal of Hydroinformatics,* vol. 25, no. 3, pp. 685-705, 2023.

[29] H. Q. Hashim and K. N. Sayl, "Detection of suitable sites for rainwater harvesting planning in an arid region using geographic information system," *Applied Geomatics,* vol. 13, pp. 235-248, 2021.

[30] S. Jaramillo, E. Graterol, and E. Pulver, "Sustainable transformation of rainfed to irrigated agriculture through water harvesting and smart crop management practices," *Frontiers in Sustainable Food Systems,* vol. 4, p. 437086, 2020.

[31] M. Abualhaija and A. H. Mohammad, "Assessing water quality of Kufranja Dam (Jordan) for drinking and irrigation: Application of the water quality index," *Journal of Ecological Engineering,* vol. 22, no. 9, pp. 159-175, 2021.

[32] M. A. Rather, G. Meraj, M. Farooq, B. A. Shiekh, P. Kumar, S. Kanga, S. K. Singh, N. Sahu and S. P. Tiwari, "Identifying the potential dam sites to avert the risk of catastrophic floods in the Jhelum Basin, Kashmir, NW Himalaya, India," *Remote Sensing,* vol. 14, no. 7, p. 1538, 2022.

[33] A. Kamel, S. O. Sulaiman, K. N. Sayl, A. A. Ali, N. S. Muhammad, J. Abdullah, and N. Al-Ansari, "The feasibility of constructing Small Hydropower Dams to maintain sustainable Water supply for domestic and agricultural uses in the Western Desert of Iraq," Research Square [Preprint], 2024.

[34] A. F. Mustafa, N. A. K. Al-Karai, and S. A. Saleh, "Comparative Assessment of Suggested Dam Sites Using both Analytic Hierarchy Process and Weighted Sum Method for Rainwater Harvesting in Al-Naft Valley, Eastern Diyala, Iraq," *The Iraqi Geological Journal,* pp. 78-98, 2024.

[35] B. Al-Hasani, M. Abdellatif, I. Carnacina, C. Harris, A. Al-Quraishi, B. F. Maaroof and S. L. Zubaidi, "Integrated geospatial approach for adaptive rainwater harvesting site selection under the impact of climate change," *Stochastic Environmental Research and Risk Assessment,* vol. 38, no. 3, pp. 1009-1033, 2024.

[36] M. N. Kasim, I. H. Obead, and I. T. Jawad, "Integrated approach for assessing suitable location of the subsurface dams in the Al Kur watershed, Iraq," *Ecological Engineering and Environmental Technology,* vol. 26, no. 2, pp. 363-378.

[37] N. Al-Ansari, S. Knutsson, S. Zakaria and M. Ezz-Aldeen. "Feasibility of using Small Dams in Water Harvesting, Northern Iraq. In: *Presented at the ICOLD Congress 2015: International Commission on Large Dams*", 2015. Available from: https://urn.kb.se/resolve?urn=urn:nbn:se:ltu:diva-38244 [Last accessed on 2025 Mar 07].

[38] B. A. Ahmad, S. G. Salar, and A. J. Shareef, "An integrated new approach for optimizing rainwater harvesting system with dams site selection in the Dewana Watershed, Kurdistan Region, Iraq," *Heliyon,* vol. 10, no. 6, 2024.

[39] Y. A. Abdulkadir, T. Baye, and M. Jothimani, "Assessing foundation characteristics at the war dam site, lake tana basin, Ethiopia: A geophysical and geotechnical perspective," *Quaternary Science Advances,* vol. 15, p. 100216, 2024.

[40] Y. Li, X. Xie, B. Jin, L. Chen, X. Liang, and K. Yin, "Comprehensive risk management of reservoir landslide-tsunami hazard chains: a case study of the Liangshuijing landslide in the Three Gorges Reservoir area," *Landslides,* vol. 22, pp. 1-21, 2024.

[41] I. Balan, C. R. Giurma-Handley, L. M. Crenganiș, F. Corduneanu, and I. E. Balan, "Slope stability analysis of a dam. Case study: Hălceni Reservoir, Iasi county, Romania," *Present Environment and Sustainable Development,* vol. 17, no. 2, pp. 141-152, 2024.

[42] V. K. Sissakian, N. Adamo, and N. Al-Ansari, "The role of geological investigations for dam siting: Mosul Dam a case study," *Geotechnical and Geological Engineering,* vol. 38, no. 2, pp. 2085-2096, 2020.

[43] B. Birhanu, Y. C. Chemeda, T. Garo, and S. Karuppannan, "Evaluation of watertightness and slope stability analysis of Upper Guder dam, West Showa, Central Ethiopia," *Quaternary Science Advances,* vol. 15, p. 100230, 2024.

[44] R. F. Hasan, M. Seyedi, and R. Alsultani, "Assessment of Haditha Dam Surface Area and Catchment Volume and Its Capacity to Mitigate Flood Risks for Sustainable Development," *Mathematical Modelling of Engineering Problems,* vol. 11, no. 7, p. 1974, 2024.

[45] M. Shahin and M. Shahin, "Runoff and River Flow," *Water Resources and Hydrometeorology of the Arab Region*, Springer, Germany, pp. 223-277, 2007.

[46] H. M. Rashid, "Reservoir sedimentation assessment using geospatial technology: A case study of Dukan Reservoir, Sulaimani Governorate, Kurdistan Region, Iraq," *Journal of Engineering,* vol. 29, no. 12, pp. 66-80, 2023.

[47] A. A. Al-Hussein, S. Khan, K. Ncibi, N. Hamdi, and Y. Hamed, "Flood analysis using HEC-RAS and HEC-HMS: A case study of Khazir River (Middle East-Northern Iraq)," *Water,* vol. 14, no. 22, p. 3779, 2022.

[48] H. El-Bagoury and A. Gad, "Integrated hydrological modeling for watershed analysis, flood prediction, and mitigation using meteorological and morphometric data, SCS-CN, HEC-HMS/RAS, and QGIS," *Water,* vol. 16, no. 2, p. 356, 2024.

[49] A. A. Othman, A. F. Al-Maamar, D. A. M. A. Al-Manmi, V. Liesenberg, S. E. Hasan, A. K. Obaid, and A. M. F. Al-Quraishi, "GIS-based modeling for selection of dam sites in the Kurdistan Region, Iraq," *ISPRS International Journal of Geo-Information,* vol. 9, no. 4, p. 244, 2020.

[50] M. L. Thieme, D. Khrystenko, S. Qin, R. E. Golden Kroner, B. Lehner,…& M. B. Mascia, "Dams and protected areas: Quantifying the spatial and temporal extent of global dam construction within protected areas," *Conservation Letters,* vol. 13, no. 4, p. e12719, 2020.

[51] A. N. Angelakıs, D. Zaccaria, J. Krasilnikoff, M. Salgot, M. Bazza, P. Roccaro, B. Jimenez,…& and E. Fereres,"Irrigation of world agricultural lands: Evolution through the millennia," *Water,* vol. 12, no. 5, p. 1285, 2020.

[52] K. N. Sayl, N. S. Muhammad, and A. El-Shafie, "Optimization of area-volume-elevation curve using GIS–SRTM method for rainwater harvesting in arid areas," *Environmental Earth Sciences,*

vol. 76, pp. 1-10, 2017.

[53] E. Umukiza, F. K. Abagale, T. Apusiga Adongo, and A. Petroselli, "Suitability Assessment and Optimization of Small Dams and Reservoirs in Northern Ghana," *Hydrology,* vol. 11, no. 10, p. 166, 2024.

[54] J. Říha, S. Kotaška, and L. Petrula, "Dam break modeling in a cascade of small earthen dams: case study of the Čižina River in the Czech Republic," *Water,* vol. 12, no. 8, p. 2309, 2020.

[55] B. H. Tessema, A. Y. Gebremedhn, and Y. S. Getahun, "Dam breach analysis and flood inundation mapping of Dire Dam, using HEC-HMS and HEC-RAS models," *Sustainable Water Resources Management,* vol. 10, no. 2, p. 45, 2024.

[56] E. Psomiadis, L. Tomanis, A. Kavvadias, K. X. Soulis, N. Charizopoulos, and S. Michas, "Potential dam breach analysis and flood wave risk assessment using HEC-RAS and remote sensing data: A multicriteria approach," *Water,* vol. 13, no. 3, p. 364, 2021.

[57] P. Bryant Robbins, "*The Evolution of Embankment Dam Design and Construction*, In: 30th Symposium of the Vancouver Geotechnical SocietyAt: Vancouver, British Columbia.

[58] S. Bhattarai, Y. Zhou, C. Zhao, and R. Yadav, "An overview on types, construction method, failure and key technical issues during construction of high dams," *Electronic Journal of Geotechnical Engineering,* vol. 21, no. 26, pp. 10415-10432, 2016.

[59] N. Adamo, N. Al-Ansari, V. Sissakian, J. Laue, and S. Knutsson, "Dam safety problems related to seepage," *Journal of Earth Sciences and Geotechnical Engineering,* vol. 10, no. 6, pp. 191-239, 2020.

[60] A. S. Toosi, E. G. Tousi, S. A. Ghassemi, A. Cheshomi, and S. Alaghmand, "A multi-criteria decision analysis approach towards efficient rainwater harvesting," *Journal of Hydrology,* vol. 582, p. 124501, 2020.

[61] M. E. Lucas-Borja, G. Piton, Y. Yu, C. Castillo, and D. A. Zema, "Check dams worldwide: Objectives, functions, effectiveness and undesired effects," *Catena,* vol. 204, p. 105390, 2021.

[62] D. W. Lee, J. S. Han, C. H. Kim, J. H. Ryu, H. S. Song, and Y. H. Lee, "Experimental and seepage analysis of gabion retaining wall structure for preventing overtopping in reservoir dams," *Applied Sciences,* vol. 14, no. 10, p. 4041, 2024.

[63] X. Fan, A. Dufresne, J. Whiteley, A. P. Yunus, S. S. Subramanian, C. A.U. Okeke, T. Pán,…&, C. T. Stefanelli, "Recent technological and methodological advances for the investigation of landslide dams," *Earth-Science Reviews,* vol. 218, p. 103646, 2021.

[64] United States Department of The Interior Bureau of Reclamation, *Design of Small Dams*. US Department of the Interior, Bureau of Reclamation, United States, 1987.

[65] N. S. Mahmood, R. Alboresha, S. O. Sulaiman, and N. A. Ansari, "Seepage problem through the foundation of a spillway with selected treatment methods," *Mathematical Modelling of Engineering Problems,* vol. 9, no. 3, 2022.

[66] N. Adamo, N. Al-Ansari, V. Sissakian, J. Laue, and S. Knutsson, "Dam safety: The question of tailings dams," *Journal of Earth Sciences and Geotechnical Engineering,* vol. 11, no. 1, pp. 1-26, 2020.

[67] A. Sidik, "Seepage handling using secant pile as a solution to replace curtain grouting and dam stability evaluation to optimize work on the Margatiga Dam construction project," *Jurnal Indonesia Sosial Teknologi,* vol. 5, no. 4, pp. 1888-1903, 2024.

[68] I. Alexandru, P. Catalin, and A. Altan, "Dam safety analysis using mathematical modelling and surveys, applied on Buftea Dam," *IOP Conference Series: Earth and Environmental Science*, vol. 1185, no. 1, p. 012017, 2023.

[69] R. Shafieiganjeh, B. Schneider-Muntau, M. Ostermann, and B. Gems, "Seepage process understanding at long-existing landslide dams through numerical analysis and hydrological measurements," *Engineering Geology,* vol. 335, p. 107524, 2024.

[70] A. M. S. Al-Janabi, A. H. Ghazali, Y. M. Ghazaw, H. A. Afan, N. Al-Ansari, and Z. M. Yaseen, "Experimental and numerical analysis for earth-fill dam seepage," *Sustainability,* vol. 12, no. 6, p. 2490, 2020.

[71] K. Narita, "Design and construction of embankment dams," *Dept. of Civil Eng., Aichi Institute of Technology,* 2000.

[72] G. Liu, Z. Zhou, J. Zhang, G. Jiang, and W. Mi, "Seepage and stability analysis of earth dams' downstream slopes, considering hysteresis in soil-water characteristic curves under reservoir water level fluctuations," *Water,* vol. 16, no. 13, p. 1811, 2024.

[73] R. Norouzi, F. Salmasi, and H. Arvanaghi, "Uplift pressure and hydraulic gradient in Sabalan Dam," *Applied Water Science,* vol. 10, no. 5, pp. 1-12, 2020.

[74] A. J. Zedan, M. R. Faris, and A. K. Bdaiwi, "Performance Assessment of Shirin Earth Dam in Iraq Under Various Operational Conditions," *Tikrit Journal of Engineering Sciences,* vol. 29, no. 2, pp. 61-74, 2022.

[75] A. Zahiri, H. Feiz Abady, and K. Ghorbani, "Estimation of sedimentation rate and storage capacity of reservoir dams using satellite imagery," *Iranian Journal of Soil and Water Research,* vol. 55, no. 7, pp. 1047-1062, 2024.

[76] I. Lawmchullova and C. U. B. Rao, "Estimation of siltation in Tuirial dam: a spatio-temporal analysis using GIS technique and bathymetry survey," *Journal of Sedimentary Environments,* vol. 9, no. 1, pp. 81-97, 2024.

[77] T. A. G. Embaye, G. H. Kahsay, N. Abadi, M. M. Kebede, and D. T. Dessie, "Evaluation of water harvesting structures on agricultural productivity: the case of Tigray Region, Ethiopia," *Sustainable Water Resources Management,* vol. 6, pp. 1-14, 2020.

[78] M. Kondolf and J. Yi. "Dam renovation to prolongreservoir life and mitigate dam impacts". Water, vol. 14,no. 9, p. 9, 2022.

[79] L. S. Othman and K. Z. Abdulrahman, "Comprehensive flow analysis of dokan dam's morning glory spillway: Utilizing advanced numerical and physical models," *Arabian Journal for Science and Engineering,* 49, pp. 1-19, 2024.

[80] R. Munoz, S. A. Vaghefi, A. Sharma, and V. Muccione, "A framework for policy assessment using exploratory modeling and analysis: An application in flood control," *Climate Risk Management,* vol. 45, p. 100635, 2024.

[81] C. Huntingford, R. Jones, C. Prudhomme, R. Lamb, J. H. Gash, and D. A. Jones, "Regional climate□model predictions of extreme rainfall for a changing climate," *Quarterly Journal of the Royal Meteorological Society,* vol. 129, no. 590, pp. 1607-1621, 2003.

[82] J. Xiong and Y. Yang, "Climate change and hydrological extremes," *Current Climate Change Reports,* vol. 11, no. 1, p. 1, 2024.

[83] M. Javansalehi and M. Shourian, "Assessing the impacts of climate change on agriculture and water systems via coupled human-hydrological modeling," *Agricultural Water Management,* vol. 300, p. 108919, 2024.

ORIGINAL RESEARCH ARTICLE UHD JOURNAL OF SCIENCE AND TECHNOLOGY

Awareness of Menstrual Abnormalities among Female Nursing Students at the University of Sulaimani

Peshwaz Abdulrahman Ahmad, Amani Fadhil Abbas, Nazera Salam Mena Qadir

Department of Maternal Neonate Nursing, College of Nursing, University of Sulaimani, Sulaymaniyah, Iraq

ABSTRACT

Background: The menstrual cycle, which occurs on a monthly basis from menarche to menopause and facilitates fertilization and conception, is a normal function in the female reproductive system. A 28-day cycle is the typical length. Any variations from the typical menstrual cycle in terms of frequency, irregularity of onset, duration of flow, or volume of blood are referred to as menstrual abnormalities. **Aim:** The current study set out to evaluate nursing students' awareness regarding menstrual abnormalities. **Materials and Methods:** In a descriptive study of the quantitative method, the sample of 100 female students was conducted at the University of Sulaimani/Nursing College from January 15 to May 30, 2024. A questionnaire format was created according to the aim of the study and delivered by a team of six experts, consisting of three parts. Part one: sociodemographic characteristics of students. Part two: Menstrual patterns of students. Part three. Awareness of students regarding menstrual abnormalities. Data were collected by direct interviews with the students. Statistical Package for the Social Science version 22 was used for analyzing the data. The frequency, percentage, and Chi-square test were used. **Results:** Results of the present study indicated that the highest percentage of participants were in the age group (20–24); they mostly dwelled in dormitory. Financial state for the majority was sufficient and the vast majority were unmarried. The majority of participants experienced painful menstruation which affected their academic performance. Moreover, only one-fifth of participants had a high awareness regarding menstrual abnormalities. Finally, the study showed that there was a significant association between the group age of students and their awareness regarding menstrual abnormalities. **Conclusion and Recommendations:** The research concludes that the majority of participants demonstrated low awareness of menstrual abnormalities. Information, education, and awareness programs need to be strengthened to spread awareness regarding menstrual abnormalities.

Index Terms: Menstruation, Menstrual Abnormalities, Nursing Students' Awareness, Dysmenorrhea

1. INTRODUCTION

The menstrual cycle is the monthly set of changes that a woman's body undergoes in preparation for pregnancy. Hormones stimulate the uterine lining to thicken with additional tissue and blood. In the event that ovulation takes place but the egg is not fertilized, the uterine lining sheds through the vaginal opening. A menstrual period is defined as the flow of blood and tissue from the uterine lining. A typical menstrual cycle might occur every 21–35 days and continue for 2 to 7 days [1].

A complicated interplay of hormones produced by the pituitary, ovaries, and hypothalamus controls the menstrual cycle. Ovulation separates the luteal and follicular phases of the menstrual cycle. Cycle length and regularity are seen as indicators of reproductive health and endocrine function [2].

Access this article online

DOI: 10.21928/uhdjst.v9n1y2025.pp65-72

E-ISSN: 2521-4217

P-ISSN: 2521-4209

Corresponding author's e-mail: Peshwaz Adulrahman Ahmad, Department of Maternal Neonate Nursing, College of Nursing, University of Sulaimani, Sulaymaniyah, Iraq. E-mail: Peshwaz.ahmad@nivsul.edu.iq

Received: 01-02-2025 Accepted: 17-03-2025 Published: 13-04-2025

A typical female's menstrual flow can last anywhere from 1 day to 8 days, although it often lasts 3–5 days. The blood loss might vary greatly, from a few spots to 80 mL, with an average of 30 mL. Blood loss exceeding 80 mL is deemed abnormal. The thickness of the endometrium, blood problems, clotting disorders, and medications are some of the variables that can impact blood flow [3].

Any variations from the typical menstrual cycle in terms of frequency, irregularity of onset, duration of flow, or volume of blood are referred to as menstrual abnormalities. Menstrual cycles are abnormal in 14–25% of women; this indicates that periods are heavier or lighter than typical, lasting longer than 35 days or <21 days. Menstrual abnormalities can happen at any age. Women under the age of 23 are the most likely to experience menstrual abnormalities, even though it is normal for women to have irregular periods at some points in their lives. Girls often experience irregular cycles for several years after menarche. Even healthy cycles in adult women can vary by a few days from month to month. Periods may occur every 3 weeks in some women and every 5 weeks in others. Flow also varies and can be heavy or light. Skipping a period and then having a heavy flow may occur; this is most likely due to missed ovulation rather than a miscarriage [1].

Menstrual disorders include disorders such as pre-menstrual syndrome (PMS), polymenorrhea, oligomenorrhea, hypomenorrhea, menorrhagia, metrorrhagia, dysmenorrhea, and amenorrhea. Menorrhagia (heavy flow), hypomenorrhea (light flow), polymenorrhea (frequent flow), oligomenorrhea (infrequent flow), and dysmenorrhea (painful flow) are common menstrual illnesses. Premenstrual syndrome, dysmenorrhea, and excessive uterine bleeding are the three most common menstrual problems in teenagers. Women may have misconceptions that what they are experiencing is normal, they may not seek medical advice unless it is directly related to pregnancy or conception [4].

Menstrual abnormalities are treated differently depending on the type of disorder and lifestyle considerations (e.g., whether a woman plans to become pregnant). Menstrual abnormalities caused by anovulatory bleeding (missing, infrequent, and irregular periods) can be treated with oral contraceptives, cyclic progestin, therapies for an underlying illness that is the source of the menstruation issue, such as nutritional therapy for eating disorders and psychotherapy. Treatment for ovulatory bleeding-related menstrual abnormalities (heavy or extended menstrual bleeding include The implantation of an intrauterine device that releases hormones and the usage of nonsteroidal anti-inflammatory drugs or different pharmaceuticals (e.g., progestin and tranexamic acid) [5].

One of the most frequent reasons women visit family doctors and gynecologists is menstrual disorders. It may have a substantial impact on women's health and result in serious health issues like osteoporosis, Type 2 diabetes, infertility, and cardiovascular disease. An abnormal menstrual cycle is associated with numerous risk factors. According to a Korean study involving 4788 women, there was a high correlation between abnormal menstruation and significant variables such as body mass index, subjective stress level, and smoking status. Furthermore, there is evidence that mental health issues such as anxiety and sadness are associated with a higher chance of abnormal menstrual cycles [6].

Menstrual disorders or abnormalities and their effects on physical and mental health are a major issue for women. More than 75% of them have some menstrual-related issues. Alterations in the typical menstrual cycles could have an impact on their overall health and mental state. Furthermore, it is commonly known that these disruptions affect social and physical activity. Concurrently, it has been discovered that menstrual issues are associated with lifestyle factors such as stress, physical activity, smoking, alcohol consumption, obesity, and exercise. In light of this, global research shows that girls are far more likely than boys to experience mental and emotional problems and behavioral issues [7].

1.1. Aim of the Study

This study was set out to determine the menstrual abnormalities experienced by female college students and their awareness. This information will be useful in modifying health promotion and education activities for young females in this environment with a view to improving reproductive health services. It is critical to encourage students about menstrual abnormalities to reduce the impact of these abnormalities on their lives, particularly their academic performance during regular sessions. Recognizing and treating menstruation disorders is an important step toward improving overall health and well-being. It is also vital that authorities focus on various programs to raise awareness for all ladies in society.

2. MATERIALS AND METHODS

2.1. Design

A descriptive quantitative study was chosen for 100 nursing girl students in the University of Sulaimani to assess student's

awareness regarding menstrual abnormalities from period January 2024 to May 2024. The data collection started on 20th January and finished on 15th February 2024.

2.2. Study Sample

A non-probability sampling approach was used in this study to choose a convenience sample of 100 female students in Nursing College/Sulaimaniyah University over a 1-month period. Those who were interested to participate were included.

2.3. Inclusion and Exclusion Criteria

For the purpose of the study, students had to fulfill the following criteria: Students who are female, enrolled in nursing programs at the College of Nursing at Sulaymaniyah and students who willingly consent to participate in the study. Students who were not enrolled in nursing programs at the designated universities, male nursing students, and those who did not give their voluntary consent to participate were not allowed to participate.

2.4. Data Tools

Direct interviews were done with study participants to gather data. After gaining agreement, a team of six trained researchers delivered the questionnaire to individuals using convenient sampling. A questionnaire format was created through a massive review of the literature. A specially designed questionnaire was used. It consists of three sections that cover sociodemographic information, menstrual pattern information, and students' awareness regarding menstrual abnormalities. The sociodemographic information included 6 questions. It consisted of information about the (age, residency, socio-economic status, mothers' education, and marital status). The menstrual pattern includes 12 questions, consisting of information about (the age of menarche, duration of flow, amount of flow, length of cycle, painful menstruation, premenstrual syndrome, experienced amenorrhea, blood clot during menstruation bleeding, dysmenorrhea interfering with academic performance, use of medications, consulted any physician, and taking treatment). The last part contained students' awareness and included 8 questions that asked for some information regarding menstrual abnormalities. A pilot research was conducted among 10 samples to test its validity. After data collection, the data were analyzed using the Statistical Package for Social Sciences (SPSS) version 22. Frequency, percentage, and Chi-square tests were used to evaluate any significant association between variables. The $P < 0.05$ is considered significant.

2.5. Validity and Reliability of the Study Instrument

The study tool for evaluating female nursing students' awareness regarding menstrual abnormality in nursing college should be subjected to expert evaluation, content, and face validity pilot testing in order to assure validity. Factor analysis can be used to demonstrate construct validity, whereas criterion validity is the process of comparing findings to pre-established measures. Test–retest methodology should be used to evaluate reliability in order to acquire a Pearson's correlation coefficient, or $r = 0.83$. Ten female students participated in a pilot study of the questionnaire, which was deemed viable. Together, these actions guarantee that the instrument assesses the intended perceptions in an accurate and consistent manner.

2.6. Data Collection

Data were gathered by conducting interviews with the sample using a specially designed questionnaire. The procedure for gathering data was carried out on 20th January 2024 to 15th February 2024. It took to 15 min for each student to be interviewed. Verbal consent was sought and the approval of the students to participate in the current study was secured. The interview took place one-on-one.

2.7. Data Analysis

The SPSS for Windows, version 22.0, was used to analyze the data. Frequency, percentage, and Chi-square tests were used to evaluate any significant association between variables. A $P = 0.05$ was used as the cutoff for statistical significance and 0.001 for high statistical significance.

3. RESULTS

Table 1 shows the sociodemographic characteristics of the study sample. 72% of the participants were aged 20–24. When it comes to student participation, grades one and two have the same percentage, representing 23% for each grade, whereas grades three and four, separately have a 27% participation rate. The majority of students (73%) are residents in the dormitory. Financial state of 89% was sufficient. Nearly half of the students' mothers were educated in primary or secondary schools, representing 45%. Regrettably, 36% of the students' mothers were not literate. 85% of the participants were unmarried, only 13% were married, and only 2% were divorced.

Table 2 demonstrates the menstrual patterns of the participants. According to the current results, 85% of students get menarche at age (12–14) years old. Regarding

TABLE 1: Sociodemographic characteristics of the study sample

Sociodemographic characteristic		Frequency	%
Age groups	<20 years	25	25.0
	20–24 years	72	72.0
	≥25 years	3	3.0
Grade	1st grade	23	23.0
	2nd grade	23	23.0
	3rd grade	27	27.0
	4th grade	27	27.0
Residency	With family	27	27.0
	Dormitory	73	73.0
Financial state	Insufficient	7	7.0
	Sufficient	89	89.0
	Highly sufficient	4	4.0
Mothers education	Illiterate	36	36.0
	Primary or secondary educated	45	45.0
	Under graduated	12	12.0
	Post graduated	7	7.0
Marital status	Unmarried	85	85.0
	Married	13	13.0
	Divorced	2	2.0
Total		100	100

TABLE 2: The menstrual patterns of the study sample

Menstrual patterns	Frequency	%
Age of menarche (years)		
≤11 years	7	7.0
12–14 years	85	85.0
≥15 years	8	8.0
Duration of flow (days)		
≤2 days	6	6.0
3–7 days	83	83.0
≥8 days	11	11.0
Amount of flow (pads/day)		
Hypominorrhea (≤2)	16	16.0
Normal flow (3–5)	75	75.0
Menorrhagia (≥6)	9	9.0
Length of cycle (days)		
≤20	36	36.0
21–35 normal	52	52.0
≥36 oligo menorrhea	12	12.0
Do you have a painful menstruation (dysmenorrhea)?		
Yes	73	73.0
No	27	27.0
Do you experience PMS (premenstrual syndrome)?		
Yes	55	55.0
No	45	45.0
Have you experience amenorrhea (absence of menstruation)?		
Yes	38	38.0
No	62	62.0
Do you have blood clots during menstruation bleeding?		
Yes	61	61.0
No	39	39.0
Does dysmenorrhea interfered with academic performance?		
Yes	48	48.0
No	52	52.0
Do you use any medication (painkiller) to relief dysmenorrhea?		
Yes	40	40.0
No	60	60.0
Have you consulted any physician on your menstrual problems?		
Yes	45	45.0
No	55	55.0
What type of treatment have you used?		
Medication	38	38.0
Surgery	3	3.0
Exercise	8	8.0
Other	6	6.0
None	45	45.0
Total	100	100

duration of flow, 83% of students had a normal duration (3–7 days) of flow. The amount of flow of 75% of participants was normal as they changed 3–5 pads per day. Half of the participants (52%) had a normal length of the cycle which was (21–35 days). Moreover, 73% of participants had dysmenorrhea which was the most common menstrual abnormality among the study samples. Dysmenorrhea affected the academic performance of 48%, making 40% of students to use a medication (painkiller) to relief dysmenorrhea. In addition, more than half (55%) of the participants had premenstrual syndrome. Nearly one-third (38%) of the students experienced amenorrhea during their reproductive life, whereas 61% experienced having blood clots with menstrual bleeding. Finally, 45% of the participated students consulted physicians on their menstrual problems, and 38% used medication as treatment.

Table 3 presents student's awareness regarding menstrual abnormalities. The vast majority (78%) had no awareness about normal timing of menarche, 74% had no awareness about normal amount of menstrual bleeding, 87% had no awareness about normal duration of normal menstruation blood flow. Unfortunately, 90% of students mentioned that severe cramps during menstruation is normal which clearly shows their poor awareness regarding this issue. The same of having blood clots with menstrual bleeding which represents 70%. In addition, 88% had no awareness that the absence of menstruation for more than 3 months is abnormal. This result supports the general observation that most students have a low-to-moderate awareness of irregular menstruation. Finally, half of the students (51%) mentioned that the severe PMS is considered abnormal.

TABLE 3: Student's awareness regarding menstrual abnormalities of the sample

Student's awareness regarding menstrual abnormalities	Frequency	%
Normal timing of menarche should be at the age of (12–13) years?		
Yes	22	22.0
No	78	78.0
Menstrual bleeding more than 6 pad changes per day is considered as heavy bleeding?		
Yes	26	26.0
No	74	74.0
Duration of normal menstruation blood flow is between 21 and 35 days?		
Yes	13	13.0
No	87	87.0
Sever cramps during menstruation is abnormal?		
Yes	10	10.0
No	90	90.0
Having blood clots with menstrual bleeding is abnormal?		
Yes	30	30.0
No	70	70.0
Experiencing severe PMS (severe premenstrual syndrome) considered abnormal?		
Yes	51	51.0
No	49	49.0
Absence of menstruation for more than 3 months is abnormal?		
Yes	12	12.0
No	88	88.0
Total	100	100

TABLE 4: Overall student's awareness regarding menstrual abnormalities

Variable	*n*=100	
	F	%
Low awareness	41	41.0
Moderate awareness	40	40.0
High awareness	19	19.0

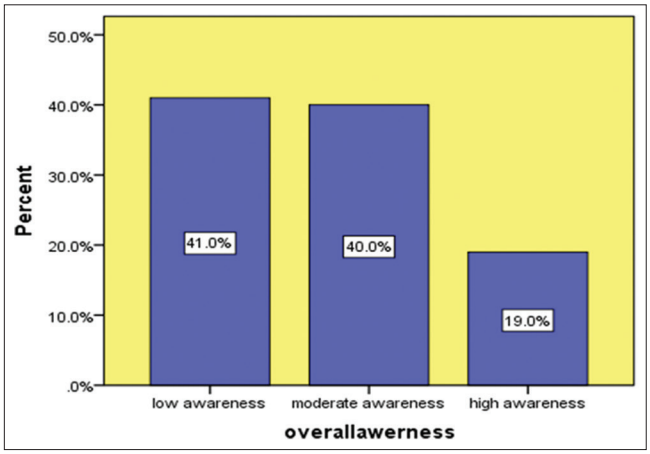


Fig. 1. Overall Student's awareness regarding menstrual abnormalities among the students.

Table 4 displays the overall awareness of students regarding menstrual abnormalities. According to the present study, 41% had a low awareness, 40% had a moderated awareness, whereas only 19% had a high awareness regarding menstrual abnormalities, see Figure 1.

Table 5 indicates the association between sociodemographic characteristics of participants and overall student's awareness regarding menstrual abnormalities. According to the results, there was no association between socio-demographic characteristics and overall student's awareness regarding menstrual abnormalities except of having a significant association between the group age of students and their awareness regarding menstrual abnormalities of $P = 0.005$.

4. DISCUSSION OF THE RESULTS

Our study aimed to explore the nursing students' awareness regarding menstrual abnormalities in the university of Sulaimani. The results support assessing students' awareness and making critical decisions to improve their awareness.

According to the presented results, the vast majority of participants were in the age group (20–24) years old, representing 72%. The group age of students in a study conducted by Manna *et al.*, 2023 [8], ranged between (17–24) years which nearly matches the present results. Regarding the grade of students, grades one and two have the same percentage of participation which represents 23% for each grade, the same is true for grades three and four which representing 27% for each grade, whereas the majority of participants in Manna and Jainendran study, 2023, belonged to 1st and 3rd years [8]. The majority of students (73%) are residents in the dormitory.

The financial state of 89% of the students was sufficient. Our results disagree with the results conducted Manna, Jainendran, 87.7% of participants belonged to the upper socioeconomic class [8]. Nearly half of the students' mothers' education was primary or secondary education which represents 45%, while unfortunately, 36% of students' mothers were illiterate. Tis results show that mothers' education plays a great role in students awareness about menstrual abnormalities, the

TABLE 5: Association between socio-demographic characteristics and overall Student's awareness regarding menstrual abnormalities among the students

Variables	*n*=100						Total	
	Low		Moderate		High			
	F	%	F	%	F	%	F	%
Age groups								
<20 years	17	68.0	6	24.0	2	8.0	25	25.0
20–24 years	22	30.6	34	47.2	16	22.2	72	72.0
≥25 years	2	66.7	0	0	1	33.3	3	3.0
P=0.005	Significant		FET=12.539				df=4	
Grade								
1st grade	16	69.6	5	21.7	2	8.7	23	23.0
2nd grade	10	43.5	9	39.1	4	17.4	23	23.0
3rd grade	8	29.6	12	44.4	7	25.9	27	27.0
4th grade	7	25.9	14	51.9	6	22.2	27	27.0
P=0.059	Not significant		x^2=12.109				df=6	
Residency								
With family	14	51.9	9	33.3	4	14.8	27	27.0
Dormitory	27	37.0	31	42.5	15	20.5	73	73.0
P=0.428	Not significant		x^2=1.814				df=2	
Financial status								
Insufficient	4	57.1	0	0	3	42.9	7	7.0
Sufficient	35	39.3	40	44.9	14	15.7	89	89.0
Highly sufficient	2	50.0	0	0	2	50.0	4	4.0
P=0.029	Not significant		x^2=10.149				df=4	
Mother education								
Illiterate	10	27.8	16	44.4	10	27.9	36	36.0
Primary or secondary	21	46.7	18	40.0	6	13.3	45	45.0
Undergraduates	6	50.0	3	25.0	3	25.0	12	12.0
Post graduate	4	57.1	3	42.9	0	0	7	7.0
P=0.307	Not significant		x^2=7.214				df=6	
Marital status								
Unmarried	35	41.2	34	40.0	16	18.8	85	85.0
Married	6	46.2	5	38.5	2	15.4	13	13.0
Divorced	0	0	1	50.0	1	50.0	2	2.0
P=0.824	Not significant		x^2=2.065				df=4	

FET: Fisher-exact-test, x^2: Chi-square, DF: Degrees of freedom

higher the mother's education higher the daughter knowledge and awreness.

Most of the participating students (85%) were unmarried, also in a study performed (Shrestha *et al.*, 2022). [9] Most participants (85.9%) were unmarried, the same results of the study conducted by Igbokwe And John-Akinola, 2021. The majority of the respondents (97.5%) were single, which matches the results of the present study.

Regarding the menstrual patterns of participants, 85% of students get menarche at age (12–14) years old. 77% of the sample study conducted by Shrestha *et al.*, 2022, became menarche at 12–14 years old which agrees with the results of our study [9]. The same results were done by Manna *et al.*, 2023, where the age of menarche for most participants was 12–14 years [8].

In the present study, 83% of students had a normal duration (3–7 days) of flow, in comparison with (Shrestha *et al.*, 2022) [9] study nearly two-thirds of samples (65.9%) had a 3–5 days duration of menstrual blood flow. The amount of blood flow of 75% of participants was normal as they changed 3–5 pads per day while participants in a study conducted by Shrestha *et al.*, 2022, changed 3–5 pads/day was of prevalence 57% [9]. Kanti *et al.*, 2020, investigated that Heavy menstrual bleeding can cause iron-deficiency anemia, which is among the leading causes of years lived with disability in low-income and middle-income countries [10].

Our study agrees with the results of Manna *et al.*, 2023 study, where about 67.4% of students had moderate blood flow [8]. Half of the participants (52%) had a normal length of cycle which was (21–35 days), while in Shrestha *et al.*, 2022 [9] study, 68.8%of participants had a normal length

of cycle which is somehow higher than the results of the present study. In a recent study conducted by Bahadori *et al.*, 2023 [11], 82.3% of participants had normal lenghth of the menstrual cycle which is higher than the present results. Moreover, (73%) of participants had dysmenorrhea which was the most common menstrual abnormality among the study samples. Dysmenorrhea affected the academic performance of 48% of them making 40% of students use a medication (painkiller) to relieve dysmenorrhea, the same results of (Manna *et al.*, 2023) [8] study, dysmenorrhea was the most prevalent gynecological issue which represented (84.0%), while in a study that conducted by Esimai and Esan, 2010, Dysmenorrhea was present in 62.5% of their study sample [12]. Quarter of the study done by Cousins and Saunders, 2023,participants missed time at college due to painful periods [13].

More than half (55%) of the participants had premenstrual syndrome, our study agrees with the results of Manna *et al.*, where the participants experienced pre-menstrual symptoms with a prevalence of 68.4% [8]. Tiranini *et al.*, similarly explored that the majority of the participants gave a history of premenstrual symptoms. According to a recent meta-analysis, premenstrual symptoms are very common, affecting about half of women of reproductive age worldwide [14].

Nearly one-third (38%) of the students experienced amenorrhea during their reproductive life, whereas 61% experienced having blood clots with menstrual bleeding. Almost half 45% of the participating students consulted a physician on their menstrual problems 38% of them used medication as treatment, whereas 16.6% of participants in a study done by Manna *et al.*, 2023, used analgesics while 45% of students did not use any type of treatment [8]. In contrast, in a study conducted by Esimai and Esan, 2010, A few students (10.5%) decided to seek help for menstrual abnormalities [12]. Kanti *et al.*, 2020, had found that dysmenorrhea was the most common menstrual abnormality. Medication was being taken mostly for dysmenorrhea [10]. The same with Bahadori *et al.*, 2023, in their study, 73.7% of students experienced a history of dysmenorrhea which made 26.2% to take medication [11]. Many individuals experience irregularities in their cycles, but due to a lack of awareness or fear of social stigma, they may not seek the necessary medical care. The physiological increase in prostaglandins during the period plays a significant role in primary dysmenorrhea. Prostaglandins, act by stimulating the contraction of uterine muscles to reduce the blood flow which is responsible for uterine hypoxia that induces painful cramping [15].

Results about Student's awareness regarding menstrual abnormalities explored that the vast majority (78%) of students had no awareness about the normal timing of menarche, 74% had no awareness about the normal amount of menstrual bleeding, 87% had no awareness about normal duration of normal menstruation blood flow. Unfortunately, 90% of students mentioned that severe cramps during menstruation are normal which clearly shows their poor awareness regarding the situation. The same of having blood clots with menstrual bleeding which represents 70%. In addition, 88% of the study sample had no awareness that the Absence of menstruation for more than 3 months is abnormal. Finally, half of the students (51%) mentioned that severe PMS is considered abnormal. According to the present study, 41% of the study sample had a low awareness, 40% had a moderated awareness, and only 19% had a high awareness regarding menstrual abnormality. Nearly, the same results of a study done by Esimai and Esan, 2010, the student's awareness of menstrual abnormalities was poor in the prevalence of (29%) [12]. On the contrary, almost all (98.9%) of the respondents were aware of menstrual disorders in a study conducted by Igbokwe and John-Akinola 2021, [16]. Mittiku 2022, conducted a study, that showed that more than one-third of the college students in Debre Berhan town have experienced menstrual irregularity. Menstrual abnormalities should not be ignored, and greater awareness is essential to improve women's reproductive health. By fostering open discussions, promoting education, and encouraging timely medical consultations, we can reduce the stigma around menstruation and ensure that all women can manage their health confidently.

According to the results, there was no association between socio-demographic characteristics and overall student awareness regarding menstrual abnormalities except for a significant association between group age of students and their awareness regarding menstrual abnormalities. This reason may be due to the absence of relationships, such as sample homogeneity or awareness complexity. The same, Esimai and Esan, performed a study, and their results showed that the awareness of students of menstrual abnormalities was significantly influenced by their age [12]. Most of the parameters did not show a significant association so a larger study or multicentric study is required.

5. CONCLUSION

In conclusion, many of the students had a low awareness regarding menstrual abnormalities, recommends the need for educational programs for college students, and secondary

and high school students in addition to parent education, work toward breaking the stigma surrounding menstruation, particularly in societies or communities where menstruation is considered a taboo topic. Schools, communities, and media can collaborate to reduce embarrassment around discussing menstruation to elevate awareness regarding menstrual abnormalities. Furthermore, there was a significant association between the group age of students and their awareness regarding menstrual abnormalities. Menstrual pain and its effect on academic performance are mentioned, which gives the study a significant new dimension. It emphasizes how menstrual irregularities have practical repercussions, particularly when discomfort prompts drug use and reduces academic performance. This feature supports the claim that menstruation health should receive greater attention in both the medical and academic spheres. Recognizing and addressing menstrual abnormalities is a crucial step toward better overall health and well-being. It is also critical to enhance authorities to focus on some programs for raising awareness for all females in society.

6. ACKNOWLEDGMENT

Gratitude is extended to all of the study participants for their participation.

7. ETHICAL CONSIDERATIONS

The scientific and ethics committees of the College of Nursing at the University of Sulaimani approved the study proposal. Before collecting data, formal authorization was obtained from health and government authorities, the study sample was provided explanation about the purpose of the study and informed consent was obtained. The participation was voluntary, and participants had right to withdraw at any time during data collection period.

8. CONFLICTS OF INTERESTS

The author affirms that they have no conflicts of interest.

9. FUNDING

The study was financially supported through author fees.

REFERENCES

[1] G. M. Attia, O. A. Alharbi and R. M. Aljohani RM, "The Impact of irregular menstruation on health: A review of the literature," *Cureus,* vol. 15, no. 11, p. e49146, 2023.

[2] S. Song, H. Choi, Y. Pang, O. Kim and H. Y. Park, "Factors associated with regularity and length of menstrual cycle: Korea Nurses' Health Study," *BMC Womens Health,* vol. 22, p. 361, 2022.

[3] D. K. Thiyagarajan, H. Basit and R. Jeanmonod, "*Physiology, Menstrual Cycle*," Statpearls Publishing, Treasure Island, FL, 2022.

[4] A. Jadwat and U. B. Bassa and Anastachia Rungusumy and Mahesh Shumsher Rughooputh, "Level of awareness on menstrual health among university students in Mauritius," *Womens Health and Wellness,* vol. 7, p. 122, 2021.

[5] Z. Bumbuliene, D. Sragyte, J. Klimasenko and E. Bumbul-Mazurek, "Abnormal uterine bleeding in adolescents: Ultrasound evaluation of uterine volume," *Gynecological Endocrinology,* vol. 35, no. 4, pp. 356-359, 2019.

[6] M. H. Alhammadi, A. M. Albogmi, M. K. Alzahrani, B. H. Shalabi, F. A. Fatta and S. F. AlBasri, "Menstrual cycle irregularity during examination among female medical students at King Abdulaziz University, Jeddah, Saudi Arabia," *BMC Womens Health,* vol. 22, no. 1, p. 367, 2022.

[7] S. Chauhan, P. Kumar, R. Patel, S. Srivastava, D. J. Simon and T. Mohammad, "Association of lifestyle factors with menstrual problems and its treatment-seeking behavior among adolescent girls," *Clinical Epidemiology and Global Health,* vol. 12, p. 100905, 2021.

[8] N. Manna, A. S. Jainendran, S. Mondal and S. Das, "A study on menstrual abnormalities among undergraduate students of medical college, Kolkata," *National Journal of Physiology, Pharmacy and Pharmacology,* vol. 13, no. 6, pp. 1246-1252, 2023.

[9] S. Shrestha, S. Manandhar, C. Limbu, C. Kunwar, P. Sinha and R. Shrestha, "Menstrual disorders and its effects on academic performance among the nursing students of PUSHS, Gothgaun," *Birat Journal of Health Sciences,* vol. 7, no. 3, p. 1871-1876, 2022.

[10] V. Kanti, V. Verma and N. P. Singh, "Study of menstrual abnormalities and its association with demographic factors among female medical students," *Journal of Clinical And Diagnostic Research,* vol. 40, pp. QC06-QC09, 2020.

[11] F. Bahadori, Z. Sahebazzamani, S. Ghasemzadeh, Z. Kousehlou, L. Zarei, M. Hoseinpour, "Menstrual Cycle disorders and their relationship with body mass index (BMI) in adolescent girls," *Journal of Obstetrics, Gynecology and Cancer Research,* vol. 8, no. 4, pp. 327-334, 2023.

[12] O. Esimai and G. Esan, "Awareness of menstrual abnormality amongst college students in urban area of Ile-Ife, Osun State, Nigeria," *Indian Journal of Community Medicine,* vol. 35, no. 1, p. 63, 2010.

[13] F. L. Cousins and P. T. K. Saunders, "Editorial: Menstruation: Myths, mechanisms, models and malfunctions," *Frontiers in Reproductive Health*, vol. 5, pp. 1-3, 2023.

[14] L. Tiranini and R. E. Nappi, "Recent advances in understanding/management of premenstrual dysphoric disorder/premenstrual syndrome," *Faculty Reviews*, vol. 11, p. 11, 2022.

[15] H. K. Al-Qazaz and R. O. Al-Dabbagh, "Menstrual disorder: Cross-sectional study on prevalence and self-care practice among adolescents in Iraq," *Annals of Tropical Medicine and Public Health,* vol. 23, no. 4, pp. 125-132, 2020.

[16] U. C. Igbokwe and Y. O. John-Akinola, "Knowledge of menstrual disorders and health seeking behaviour among female undergraduate students of University of Ibadan, Nigeria," *Annals of Ibadan Postgraduate Medicine,* vol. 19, no. 1, pp. 40-48, 2021.

REVIEW ARTICLE

UHD JOURNAL OF SCIENCE AND TECHNOLOGY

Deep Learning Approaches for Retinal Disease Identification in Fundus Imaging: A Comprehensive Overview

Ismael Abdulkareem Ali, Sozan Abdullah Mahmood

Department of Computer, College of Science, University of Sulaimani, Sulaymaniyah 46001, Kurdistan, Iraq

ABSTRACT

Vision impairment is becoming a major health concern, especially in elderly people. While in the medical field, manually detecting ocular pathology has significant difficulty. Therefore, deep learning diagnostic techniques are widely used for identifying eye diseases and play a crucial role in diagnosing vision-related problems. Examination of fundoscopy allows for analyzing eyes for diagnosis of eye diseases, including Diabetic retinopathy (DR), Cataracts, Glaucoma, Age-related macular degeneration, Pathologic Myopia, and more. In this paper, we propose a concise review of introducing most of the DL, hybrid, and ensemble models utilized for the purpose of identification and classification of eye diseases. Various datasets, feature extraction techniques, and metrics for performance evaluation are discussed. The chosen papers come from conferences and scholarly publications published from 2019 to 2024. We evaluate the performance of chosen researches using different datasets, the most common ones include ocular disease intelligent recognition, Indian DR image dataset, EyePACS, methods to evaluate segmentation and indexing techniques in the field of retinal ophthalmology, methods to evaluate segmentation and indexing techniques in the field of retinal ophthalmology-version 2, DIARETDB, Structured analysis of the retina, high-resolution fundus, digital retinal images for vessel extraction, online retinal fundus image dataset for GI analysis and research, retinal fundus multi-disease image dataset and Kaggle datasets. The detection studies that have been reviewed show that the accuracy of these approaches varies between 73% and 99%, the sensitivity ranges from 69% to 99% and precision is between 89% and 99%. The results show that great accuracy is consistently achieved with DL algorithms compared to traditional Machine learning approaches. However, there are still some challenges and limitations remaining including excessive resource consumption and over-fitting due to dataset size and diversity issues. This review offers useful insight for researchers and healthcare professionals to comprehend AI technologies properly for the detection, classification, and diagnosis of retinal diseases. We succinctly summarize the methodologies of all the chosen studies and focus on the elements that define the aim of the studies.

Index Terms: Eye Diseases, Retinal Disease Diagnosis, Color Fundus Mages, Hybrid Deep Learning, Deep Learning

1. INTRODUCTION

The eyes are considered to be one of the most important sense organs for daily functioning and crucial for observing and perceiving the outside world. Eyesight problems are experienced by almost everyone at some point in their lives. Over the years, ophthalmologists and healthcare centers always trying to find treatments and cures for people with retinal disorders that cause blindness. The World Health Organization reported in August 2023 that at least 2.2 billion individuals suffer from a sightedness problem. They state that all ages can be affected by vision loss, but most individuals who experience vision impairment and blindness are over 50 [1]. In 8 out of 10 cases, blindness can be prevented or

Access this article online

DOI:10.21928/uhdjst.v9n1y2025.pp73-92

E-ISSN: 2521-4217

P-ISSN: 2521-4209

Corresponding author's e-mail: Sozan Abdullah Mahmood, Department of Computer, College of Science, University of Sulaimani, Sulaymaniyah 46001, Kurdistan, Iraq. E-mail: sozan.mahmood@univsul.edu.iq

Received: 14-01-2025 Accepted: 25-02-2025 Published: 19-04-2025

avoided with early intervention or suitable treatment, but before providing an appropriate treatment for a patient's health problem getting the right identification and disease diagnosis is a key aspect of healthcare problems.

A particular type of machine learning known as "deep learning" (DL), which involves training multi-layered artificial neural networks to find patterns in data. The development and improvement of highly effective DL techniques have grown in healthcare as a promising tool for disease detection and classification, Drug identification and recognition, Medical image analysis and diagnosis, image segmentation, and recognize particular anatomical features or lesions. Thus, to improve the quality of treatment and health for patients, physicians, and other medical professionals, DL is being implemented in the present healthcare system [2].

Obtaining a retinal image using the appropriate medical image modality is the initial stage in the process of detecting and diagnosing retinal disorders [3]. With the invention of various imaging modalities over the years, "Retinal or fundus photography" has become more well-known because it is a fast and straightforward procedure while also considerably cost-effective and non-invasive. It is capable of capturing the retina, blood vessels, optic disc (OD)/cup (OC), macula, fovea, and posterior pole on the interior surface of the eye [4], see Fig. 1. Fundus photography images can hold and indicate some serious abnormities in diseased eyes including Diabetic retinopathy (DR), glaucoma (Gl), cataract (CA), Age-related macular degeneration (AMD), hypertension, Pathologic Myopia, and other diseases, which are the primary worldwide causes of blindness and visual impairment. Fig. 2. shows fundus images of various retinal disorders.

Diabetes patients may experience visual loss and blindness due to a condition called DR. Certain lesions in the eyes are caused by DR, and over time, these lesions can lead to permanent blindness. Lesions include microaneurysms (MAs), hemorrhages (HAs), exudates (EXs), cotton wool spots, and abnormalities in retinal blood vessels [6]. The International Agency for the Prevention of Blindness reports one in three diabetic people have some kind of DR, and one in ten will experience a vision-threatening condition. DR has 4 stages, Mild, Moderate, Severe non-proliferative retinopathy (NPDR), and Proliferative DR (PDR). MAs, or tiny bulges in the blood vessels of the retina, can leak fluid into the retina in the early stages of the disease (Mild). In addition, this may result in macula swelling, which impairs vision. At this point, symptoms are typically minimal or non-existent. During the Moderate phase, the retina's blood vessels enlarge and may block. This may exacerbate diabetic macular edema (DME), a condition in which fluid accumulates in the macula area of the retina and impairs or destroys vision. The number of blocked blood vessels in the eye is growing in Severe NPDR. Retinal growth of new blood vessels is subsequently signaled. One may experience fuzzy vision with dark spots if the blood vessels totally close off. The retina loses oxygen in the last stage (PDR), and new blood vessels develop inside the retina and into the vitreous gel, which fills the eye. These blood vessels are fragile and could burst and hemorrhage [7].

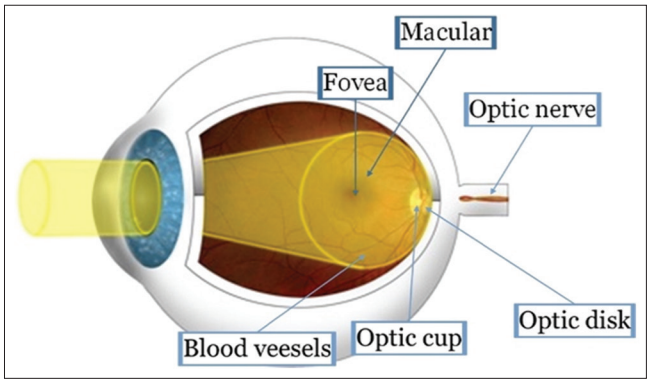


Fig. 1. Color fundus photography [5].

Another common eye disease is Gl, which is caused by damage to the optic nerve that connects the eye to the brain. Usually, it results from fluid accumulation in the anterior chamber of the eye, which raises intraocular pressure. Although it can afflict individuals of any age, folks in their 70s and 80s tend to be the most affected [8].

A CA is a disorder in which the normal lens of the eye gets clouded. When the proteins in your lens degrade, objects appear foggy, fuzzy, or less vibrant. If someone has this eye condition, their vision may shift to the point where they see bright colors as faded or yellow instead, they may see double or ghostly images out of their eyes. Experiencing increased light sensitivity, and/or having poor night vision. It is necessary to inform your ophthalmologist if you experience any of these signs of a CA [9].

Another typical condition affecting the eye's focus is known as nearsightedness or myopia when far objects appear blurry while close-up items appear clear. For example, you can see clearly enough to read a map, but not well enough to drive a car. For several decades, the prevalence of myopia has been rising and nearly half of the world's population is predicted to be nearsighted by 2050. Other indications of myopia include headaches, eyestrain, and squinting to see clearly [10].

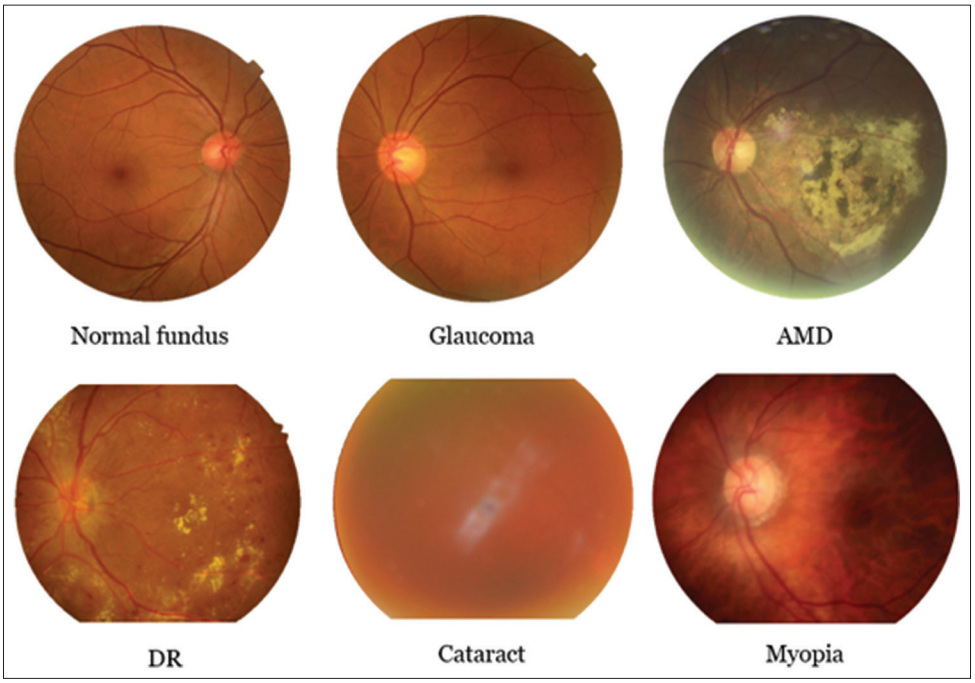


Fig. 2. Fundus images with different eye conditions.

AMD is a retinal issue. AMD is quite prevalent and it is one of the main causes of visual loss in those over 50. It occurs when there is damage to the macula (a portion of the retina). Loss of center vision is a symptom of AMD. However, your side vision, or peripheral vision, will remain normal. Whether you are looking at something close up or far away, you cannot see fine details. Consider the scenario where you are staring at a clock that has hands. While you have AMD you may be able to see the numbers around the edge of the clock clearly, but not the hands [11].

This paper aims to offer a systematic review to shed light on the advancements made in DL, hybrid, and ensemble-based techniques that are especially intended to recognize, classify, and grade retinal conditions. By focusing on recent research publications, we critically evaluate existing methodologies, highlight their advantages and limitations, and identify emerging trends. The ultimate goal is to provide insights that guide future research toward developing more accurate, explainable, and clinically viable AI-based diagnostic tools.

2. METHODS

2.1. Retinal Image Databases

In this section, we provide information about the datasets used for retinal disease diagnosis. The datasets used by the studies comprise private and publicly accessible ones. Table 1 shows the summary of fundus image datasets for diagnosing retinal disorders

2.1.1. Indian DR image dataset (IDRiD) [12]

It is a publicly available retinal fundus image database that provides information on disease severity grading of DR and DME. Furthermore, provides different retinal lesions associated with DR. Including MAs, Hemorrhages (HEs), and Soft & Hard EXs. It consists of a total of 516 images, which contain images marked with DR and/or DME and 134 Normal ones (without signs of DR and/or DME).

2.1.2. High-resolution fundus (HRF) image database [13]

The dataset is public and freely available for research purposes. Presently, there are 15 images in the database showing healthy ones, 15 of them showing patients with DR, and the other 15 images showing patients with Gl. For every image, binary gold standard vessel segmentation images are supplied Budai and Odstrcilik [14].

2.1.3. Methods to evaluate segmentation and indexing techniques in the field of retinal ophthalmology (Messidor) [15]

The purpose of the Messidor database is to support research on computer-aided diagnosis of DR. The French Ministry of Research and Defense supported the Messidor

Table 1: Summary of fundus image datasets for diagnosing retinal disorders

Dataset	Samples	Image distribution pre-class	Purpose	Disease diagnosis	Studies
DRIVE [17]	40	N: 33|Mild DR: 7	Blood vessel segmentation	DR	[38]
IDRiD [12]	516	N: 134|Mild DR: 20, Moderate DR: 136|severe DR: 74|PDR: 49	Detection, Segmentation, and grading	DR and DME	[38]–[42]
RFMiD [19]	3200	N: 669|DR: 632|ARMD: 169|MYA: 167|Other 43 categories: rest	Multi-disease Classification	46 conditions	[41], [43], [44]
ODIR [21], [22]	6426	N: 3098|DR: 1406|GL: 224|CA: 265|AMD: 293|H: 107|MYA: 242|other diseases: 791	Multi-Label multi-class Classification	8 conditions	[39], [41], [45] –[50]
Messidor [15]	1200	-	Exudates Detection, DR Grading, and OD/OC segmentation	DR and DME	[37], [39], [40], [50]–[52]
Messidor-2 [15], [23]	1748	RDR: 455|NRDR: 1286	DR diagnosis and lesion segmentation	DR and DME	[39], [42], [51], [53]
STARE [54]	20	N: 10|non-N: 10	Blood vessel segmentation	-	[38], [40], [55]
HRF [13]	45	N: 15|DR: 15|GL: 15	DR diagnosis and Blood vessel segmentation	DR and Glaucoma	[38], [41]
ORIGA [24]	650	N: 482|GL: 168	Glaucoma diagnosis and OD/OC segmentation	Glaucoma	[40], [56], [57]
REFUGE [25]	1200	N: 1080|GL: 120	Glaucoma classification and OD/OC segmentation	Glaucoma	[57]
DIARETDB0 [27]	130	N: 20|DR: 110	MAs, HMs, H&SO EXs segmentation	DR	[50]
DIARETDB1 [29]	89	N: 5|Mild NPDR: 84	DR grading	DR	[42], [50]
ARIA [17]	450	-	ONH boundary Segmentation	AMD and DR	[40]
DRISHTI-GS [30]	101	-	ONH segmentation	Glaucoma	[39], [51], [58]
DRISHTI-GS1 [32]	101	N: 31|GL: 70	ONH segmentation	Glaucoma	[56]
Kaggle (EyePACS) [33]	88,702	-	DR grading	DR	[37], [59], [60]
Kaggle (APTOS-2019) [34]	3662	-	DR grading	DR	[61]–[63]
Kaggle (eye disease classification) [35]	4217	N: 1074|CA: 1038|DR: 1098|GL: 1007	Multi-class Classification	DR, GL, and CA	[36], [64]

DRIVE: Digital retinal images for vessel extraction, IDRiD: Indian DR image dataset, HRF: High-resolution fundus, STARE: Structured analysis of the retina, RFMiD: Retinal fundus multi-disease image dataset, ODIR: Ocular disease intelligent recognition, ORIGA: Online retinal fundus image dataset for GI analysis and research, REFUGE: Retinal fundus GI challenge, DIARETDB0: Standard DR database calibration level 0, DIARETDB1: Standard DR database calibration level 1, ARIA: Automatic retinal image analysis, APTOS-2019: Asia Pacific Tele-Ophthalmology Society 2019, DR: Diabetic retinopathy, ONH: Optic nerve head, AMD: Age-related macular degeneration, EXs: exudates, MAs: microaneurysms, PDR: Proliferative DR, NPDR: retinopathy (NPDR)

research program as part of the 2004 TECHNO-VISION program. This database includes 1200 TIFF images with an Excel file containing each image's medical diagnosis. ADCIS [16].

2.1.4. Digital retinal images for vessel extraction (DRIVE) [17]

DRIVE dataset has been created to facilitate comparative research on blood vessel segmentation in retinal pictures. The images used in the DRIVE database were from a Dutch study that screened for DR. There are two sets of 40 photographs total: a training set and a test set, each with 20 images. 7 exhibit mild early DR and 33 do not exhibit any indications of the condition.

2.1.5. Structured analysis of the retina (STARE) [18]

Stare is a dataset where its challenge is in the retinal vessel segmentation and this is used in medical analysis. Out of the total twenty fundus color retinal images in the STARE dataset, ten are healthy ocular fundus images and the other ten ocular fundus images are unhealthy. The dataset contains two sets of hand-annotated images segmented by two human experts as ground truth for retinal blood vessel segmentation methods. STARE: A Retinal Image Dataset [18].

2.1.6. Retinal fundus multi-disease image dataset (RFMiD) [19]

A retinal image collection called RFMiD is made available to the public. It includes 3200 photos annotated by experts

that were taken with three distinct fundus cameras. Three subsets of the dataset are created: 20% for evaluation (640 photos), 20% for testing (640 images), and 60% for training (1920 images). Dataport [20]

2.1.7. Ocular disease intelligent recognition (ODIR) [21], [22]

ODIR-5K is a structured ophthalmic database of 5,000 volunteer patients along with age, color fundus images (CFI) of their left and right eyes, and diagnostic keywords performed by expert ophthalmologists. The images are saved in JPG format with various sizes and dpi values. The diseases are labeled in a CSV file with only one letter including Normal (N), CA (C), Gl (G), Diabetes (D), AMD (A), Pathological Myopia (M), Hypertension (H) and Other diseases/abnormalities (O)

2.1.8. Methods to evaluate segmentation and indexing techniques in the field of retinal ophthalmology-version 2 (Messidor-2) [15], [23]

The Messidor-2 dataset is a set of 874 DR examinations (1748 images), with two macula-centered eye fundus images one per eye. It is free to use for academic and research purposes and is not affiliated with any commercial interest. The dataset also comes with a spreadsheet showing the pairing of images. Annotations like a "ground truth' for DR are missing in them. Messidor-2 can be used, free of charge, for research and educational purposes at Messidor2: A Retinopathy Dataset [23].

2.1.9. Online retinal fundus image dataset for Gl analysis and research (ORIGA) [24]

This dataset incorporates 650 retinal photographs annotated by qualified experts from the Singapore Eye Research Institute. It was created to diagnose Gl and to segment the OD and OC. For every image presented the CDR value, the OD and OC boundaries, and a label identifying the presence or absence of Gl. The public can access ORIGA upon request.

2.1.10. Retinal fundus Gl challenge (REFUGE) [25].

An openly available data set of 1200 fundus photographs with clinical Gl labels and ground truth segmentations is provided by REFUGE. Online available at Orlando *et al.* [26].

2.1.11. Standard DR database calibration level 0 (DIARETDB0) [27]

This database includes 130 retinal images that were captured with a 50° field of view digital fundus camera. About 20 of the 130 photos are in good health, and the other 120 show signs of DR. Each image is labeled with the presence or absence of red tiny dots, hemorrhages, hard EXs, soft EXs, and neovascularization. Available through the webpage at DIARETDB0 [28].

2.1.12. Standard DR database calibration level 1 (DIARETDB1) [29]

Retinal fundus images in this database are 89. 84 of them show mild NPDR, and the remaining 5 photos are normal. The photographs were taken with a camera with a 50° field of view. Both the DIARETDB1 and DIARETDB0 databases provide images with a resolution of 1500 × 1152.

2.1.13. Automatic retinal image analysis (ARIA) [17]

There are 450 images in the JPEG format ARIA database and they were annotated by two expert ophthalmologists. Three groups are created from those images: one with AMD, one with DR, and one with a healthy control group.

2.1.14. DRISHTI-GS: Retinal image dataset for optic nerve head segmentation [30]

A dataset of retinal images for OD and OC segmentation. A total of 101 photos make up the DRISHTI-GS collection. 50 training and 51 testing photos make up the set. With the patients' permission, every photograph was taken at the Aravind Eye Hospital in Madurai. The dataset is available at Drishti-GS: Glaucoma Dataset [31].

2.1.15. DRISHTI-GS1: Retinal image dataset for optic nerve head segmentation [32]

This dataset is an extension of DRISHTI-GS. It contains 50 and 51 training and testing images, respectively. Manual segmentations are obtained for both the OD and OC regions of each image from 4 different human specialists with differing levels of clinical experience. A dataset was recently made publicly available at Drishti-GS: Glaucoma Dataset [31].

2.1.16. Kaggle (EyePACS): Kaggle dataset for DR [33]

The dataset consists of 88,702 high-definition retinal photos obtained in diverse imaging conditions. Each image has a clinician's rating, from 0 to 4, indicating the presence of DR. This Kaggle dataset consists of 35,126 training samples and 53,576 test samples in total. It is available on Kaggle [33].

2.1.17. Kaggle: The Asia Pacific Tele-Ophthalmology Society 2019 (APTOS-2019) [34]

The Kaggle dataset has 3662 samples gathered from numerous people in rural India and then it was arranged by the Aravind Eye Hospital. The fundus images were taken over an extended period in various settings and situations. There

are five categories for the samples based on DR Disease Severity. It is available on Kaggle [34].

2.1.18. Kaggle (eye_diseases_classification) [35]

The dataset includes more than 1000 retinal images per class, including those for normal, DR, CAs, and Gl. Totally, this dataset includes 4217 fundus images of three different kinds of eye diseases and a normal class as well. The sources of such images are diverse and include IDRiD, Ocular recognition, HRF, and others. It is available on Kaggle [35].

2.2. Pre-processing Stage

Images typically vary in brightness, intensity, and visual quality. In this section, we discuss pre-processing methods frequently employed for fundus image analysis. Gaining the data in a structure and shape, while also cleaning raw data to build and train Convolutional Neural Network (CNN) and machine learning models. Attaining greater accuracy is a crucial stage in this process. Pre-processing retinal pictures is typically necessary for the prediction of eye diseases to increase model performance and accuracy. Table 2 provides an overview of various pre-processing methods used in ocular disease diagnosis studies.

Pre-processing the retinal photographs to extract the Region of Interest (ROI) is the first step in Kadum *et al.* [36]. After that, they utilized suitable image-processing methods to separate the pertinent regions associated with every individual ocular condition. This reduced the impact of background noise and unnecessary image areas to concentrate the analysis on the areas with abnormalities related to the disease. After the ROI extraction, we scaled the pictures to 512 × 512 pixels, which is a common size. This process guaranteed uniformity in the input data and enabled effective feature extraction.

Developing a hybrid image enhancement algorithm to improve the quality of pictures is proposed in the study [37]. Thus increasing contrast and reducing noise to support the standard of color fundus imaging. There are two main phases to the approach: First chopping images to remove unnecessary information. Then the use of Gaussian blurring and cropping to improve contrast and reduce noise. After receiving retinal fundus photographs, the Fundus Image Enhancement (FGB) model identifies the foreground and uses Gaussian blurring to highlight the retinal vessels. Followed by feature extraction and classification.

This dataset [65] is used by Londhe [66] consisting of (300) normal and 3 other categories (100/category) of eye diseases. Images come in a variety of dimensions. The pictures were resized to the same dimensions. Then, the black background of the images is removed with cropping and resized again to 224 × 224. Then, the augmentation step is performed (horizontally and vertically flipping) to ensure that each class has 300 images. The dataset is split into 70%, 10%, and 20% for training, testing, and validation, respectively. Normalization as a data transformation step is utilized next, in which 8-bit channels of the colors red, green, and blue are used to represent images.

In hybrid [53] the dataset used is Decencière *et al.* [15], Messidor2: A Retinopathy Dataset [23], which consists of variously sized photos. Thus, their resolution changed to 960w × 1200h. After that normalization operation is performed before providing images as input to a DL model, each image's pixel values should be scaled, from the range of [0–255] to the of [0–1] by dividing them to 255. Finally, the feature extraction procedure receives pre-processed images.

The pre-processing algorithm proposed in Xu *et al.* [62]. Starting by cropping the image's non-retinal portion. Gaussian Blur method is used for enhancement of images which improves contrast and reduces noise. For comparative purposes in later experiments, a colored version of the clipped image is additionally processed. Automated and Center cropping were also carried out. Data augmentation is finally applied to the pre-processed dataset [34] to prevent over-fitting and expand the dataset's diversity.

The first step in Menaouer *et al.* [63] for pre-processing was reading images and resizing them to 128 W ×128 H. To enhance the CNN models' training performance, the pixels are normalized. Then, to reduce overfitting, real-time image augmentation was used, and to distribute the data evenly among the severity levels of DR. After augmentation, 5000 photos were obtained for each class. Finally, hybrid DL approaches used the generated images as input.

In Mahmoud *et al.* [67] a system was employed to optimize the brightness, illumination, and equalization of the photographs. The pre-processing step involves enhancing the contrast and brightness of the input photos and removing any unwanted background details. To enhance image clarity, colorful fundus photos are normalized to a specific luminance level.

For this study [57], pre-processing was carried out in two stages, (1) RGB2GRAY: authors converted RGB images to grayscale because Retinal images in grayscale mode are better than RGB at extracting and identifying textural features, additionally lowering the noise and improving the outcome.

Table 2: Pre-processing methods used in chosen studies

No.	Studies	Pre-processing methods	Strengths	Weaknesses
1	Doddi [36]	• ROI extraction • Resizing	• Focuses on relevant regions that include crucial information, standardizes the size to ensure consistency across input data	• This may result in the loss of peripheral details and useful information
2	Abbood *et al.* [37]	• Greyscale image cropping • Circle crop and GaussianBlur	• Performs foreground identification of fundus images and applies Gaussian blurring to enhance the visibility of retinal vessels	• Circular cropping may remove valuable information
3	Retina_ Dataset [66]	• Cropping • Resizing • Data augmentation (flip vertically & horizontally) • Normalization	• Eliminates redundant regions, increases dataset diversity, enhances model generalization, and accelerates the training process	• Data augmentation may introduce unrealistic transformations
4	Khan *et al.* [53]	• Resizing • Normalization	• Standardizes input images for CNNs and preferred in DNN model training	• Not applying de-noising techniques or contrast enhancement (e.g., CLAHE)
5	Butt *et al.* [62]	• Cropping • GaussianBlur pre-processing (image enhancement) • Color version of cropping & GaussianBlur's pre-processing • Auto cropping • Center cropping	• Multiple enhancement techniques improve adaptability, by automatically cropping the ROI based on the retinal image's properties; the automatic cropping algorithm can increase pre-processing's accuracy and efficiency.	• Certain techniques may distort important features and destroy the color information of the image
6	Xu *et al.* [63]	• Resizing • Normalization • Augmentation • Dataset splitting into 50:25:25 (Train, Test, and Val)	• These steps enhance processing efficiency, improve model performance, increase dataset diversity, and ensure reliable evaluation	• The model may struggle with noise and changes in image quality because the authors didn't introduce techniques to reduce noise or improve contrast in their pre-processing steps
7	Londhe [67]	• Contrast enhancement	• Images often differ in intensity, brightness, and quality, Standardizing these factors helps improve the performance of the ML algorithm	• May amplify noise
8	Thanki [57]	• RGB2GRAY • Texture feature extraction (by Gray-Level Co-Occurrence Matrix)	• Reduces computational complexity as grayscale images have Fewer Channels to Process, extracts important texture features	• Loss of color information and reduce models performance for tasks that depend on color features
9	Nawaldgi and Lalitha [59]	• Image resizing • Data augmentation • Applying a median filter • Image sharpening	• Enhances image quality, reduces noise	• Sharpening can introduce noise or unwanted artifacts
10	Verma *et al.* [61]	• Image resizing • Normalization	• Prepares images to meet the input requirements for the GoogleNet and ResNet-18 models	• No explicit enhancement techniques are described
11	Ouda *et al.* [44]	• Up-sampling • Cropping • Resizing • contrast	• The dataset had an uneven distribution of images across disease classes, upsampling was performed to help the model learn more effectively	• While upsampling helps balance the dataset, it can lead to overfitting by repeating minority class images, causing the model to memorize patterns instead of learning to generalize
12	Mahmoud *et al.* [68]	• Resizing to 224×224 • Normalization • Dataset splitting into 80:20 (Train, Test)	• Standard input size for CNNs ensures proper dataset division	• No additional feature enhancement
13	Ishtiaq *et al.* [60]	• Remove duplicates • Remove flow identifiers • Feature conversion • Dataset splitting • Feature normalization • Replace infinite values to 0	• Cleans data for better model performance	• May discard potentially useful data
14	Shimpi and Shanmugam [69]	• Resizing • label encoding • Data augmentation	• Standardizes input improves class balance	• No explicit enhancement techniques are described

(Contd...)

Table 2: *(Continued)*

No.	Studies	Pre-processing methods	Strengths	Weaknesses
15	Sarode and Desai [70]	• Resizing • Image enhancement (CLAHE) • Dataset splitting • Data augmentation • Normalization • Resizing to 496×496	• CLAHE enhances contrast effectively for medical images	• Computationally intensive. If parameters are not properly tuned, CLAHE can produce suboptimal results
16	Menaouer *et al.* [64]	• Average filters • Laplacian filters • Data augmentation • Dataset splitting into 80:20 (Train and Test)	• Enhances edge details, improves contrast	• Filtering may remove subtle features

CLAHE: Contrast limited adaptive histogram equalization, CNN: Convolutional neural network

(2) Feature Extraction: The Gray-Level Co-Occurrence Matrix (GLCM) approach is used to extract texture information from retinal fundus pictures.

Preprocessing the dataset was the initial stage of the suggested model [59]. To improve the quality of images, enhancement techniques are used. First, dataset images were resized (to 512 × 512) to standardize them because they have different dimensions. Then, a data augmentation method was employed to achieve data balance. To remove noise from the images, a median filter was applied to the full dataset. Finally, The sharpening filter operated by first creating a blurry copy of the original image, which was then subtracted from the original.

In addition to significant imbalance, the data also have issues with focus, exposure, noise, and artifacts. Pre-processing involved in this study [61] was resizing and normalizing the photographs to comply with GoogleNet and ResNet-18 Models' specifications for input images (224 × 224 × 3).

The authors of Abbas *et al.* [44] used a range of pre-processing methods to provide a diversity of training data. The primary method involved was upsampling the dataset for the model to learn better. After that, the dataset's images were cropped to have a 1:1 ratio size so they could be quickly and easily resized to 224 × 224. The dataset's original images included a range of resolutions. This process of cropping enables keeping the essential information to improve learning during model training.

The authors in this research [68] pre-process the Kaggle dataset by resizing images to the standard size of 224×224 pixels and performing normalization. Data was split (80/20) for training and testing was also performed.

Authors in Verma *et al.* [60] mention that data-cleaning methods such as de-noising, normalizing, and equalization can be applied to fundus images [33] to enhance image quality, as these images may include artifacts such as reflection, noise, and brightness variation. As further data augmentation approaches, flipping, rotating, and zooming can be used to reduce overfitting and diversify datasets. Pre-processing techniques can be used on the images to minimize dimensionality and enhance the effectiveness of the feature extraction and ensemble learning phases once the dataset has been cleaned, augmented, split, and labeled.

After gathering the retinal fundus images from Emma Dugas and Cukierski [33], pre-processing steps, such as scaling, cropping, and "Contrast Limited Adaptive Histogram Equalization (CLAHE)" are applied in Ali Tabtaba and Ata [42]. In the image augmentation phase, a high-quality image is provided for additional processing using Generative Adversarial Networks (GANs). Finally, the HCMD-CNN model takes the augmented images for subsequent processes. The pre-processing techniques applied in Sarode and Desai [69] were made up of 3 steps. (1) Resizing – the technique that is most commonly used for resizing to change the dimensions of images is a bilinear interpolation. (2) A technique called label encoding is applied to translate class or categorical labels into numerical representations. (3) Data augmentation – an effective pre-processing method that artificially increases the training dataset's variety. Several transformations are utilized including Rotation, Width and Height Shift, Shear Range, Zoom Range, and Horizontal Flip.

In Pin *et al.* [70], image pre-processing methods are applied before feeding OCT images to neural network models, such as resizing images to 500 × 300. After that CLAHE technique was performed to enhance the quality of images (clip limit = 2.0 and tile grid size = 3×3). There are three

sets in the dataset Division. First: The dataset is randomly divided into 80% train and 20% test sets. Second K-Fold cross-validation is applied to separate the train set into five holes for training and one for validation. Finally, three sets are created from the dataset. Data augmentation was used to increase dataset size, prevent overfitting, and generalize image classification models.

This study [64] applied enhancing filters to CFP images to improve their quality and highlight the boundaries of the regions of interest. An average filter is utilized to enhance CFP images from a dataset, smoothing them and reducing noise. The Laplacian filter sharpens border delineation, making it easier to identify edges and features in the images. By employing various data augmentation techniques, including flipping, rotating, shifting, and more, the number of images during the training phase is increased. Finally, the dataset [35] is divided with an 80:20 distribution for training and testing. Table 2 lists the pre-processing methods utilized by hybrid models.

3. HYBRID DL

This section describes how the research studies were conducted. According to Bouchrika [71] the research paper's methodology is the foundation for evaluating the validity of your investigation. In this section, we discuss the methodology followed by each study chosen, and almost all of them utilize hybrid techniques for their work. Fig. 3. shows the distribution of reviewed papers published from 2019 to 2024.

This study [66] introduces a novel method for classifying 4 different categories of eye diseases using the benefits of combining CNN-RNN models (hybrid). The early stages of the methodology involve data collection from Kaggle, exploratory data analysis, pre-processing, and data transformation. In the modeling phase, CNN would serve as a feature extractor which is being carried out using the chosen transfer learning models, InceptionV3, DenseNet169, and InceptionResNetV2. The collected features are then classified using the LSTM module, which is an RNN model. The model was implemented on both imbalanced and balanced datasets. The pre-trained weights of each of the six constructed models are used in the transfer learning models. K-fold cross-validation is used to implement each of the six models at k=5.

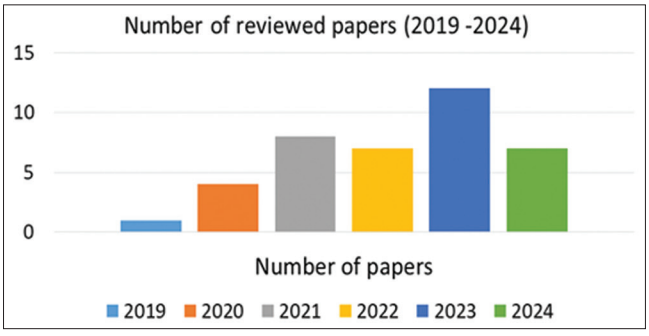


Fig. 3. Distribution of reviewed papers published per year (2019–2024).

To improve overall image quality and contrast, this paper [37] presents a hybrid image enhancement model designed specifically for CFI and five phases of early diagnosis of DR (normal, mild, moderate, severe, and PDR). The Kaggle dataset (EyePACS) and MESSIDOR were used in this article. After receiving retinal images, the model uses a modified Resnet50 network for DL feature extraction and DR classification, while also foreground detection and Gaussian blurring are used to emphasize retina vessels.

A model for multi-retinal disease classification (MRDCM) was created in this study [45]. This model utilizes a combination of ensemble, transfer, and DL techniques, focusing on the top four diseases. The process began with selecting the dataset, determining the optimal DL model, and applying data augmentation and preprocessing methods. It then employed advanced training techniques and implemented an ensemble of EfficientNetB4 and EfficientNetV2S to create the final model. The subsequent step involved interpreting the model's predictions and diagnosing images by assessing the proportion of each disease within them. For individual classifications, the researchers ultimately stacked binary logistic regression models.

The authors in Shimpi and Shanmugam [68] presented a novel hybrid model (DRNN) for DR detection and grading. It comprises two distinct innovative DL architectures, ResNet-152 and DenseNet-121. Acquiring the DR Kaggle dataset, pre-processing it by normalizing and resizing the images, dividing the training and test data in 80/20% and the applying proposed hybrid DRNN model with some hyper-parameters are the main steps of the methodology for this work. The DRNN model is constructed using many ResNet-152 layers, then multiple DenseNet-121 levels. Each layer has numerous convolution and pooling layers. The DenseNet-121 blocks extract more detailed features from the input images and the ResNet-152 blocks extract higher-level information from it. To achieve the predicted class labels, a classification layer is inserted at the network's end.

The pre-trained CNN models of GoogleNet [72] and ResNet-18 [73] are used in [61] to create a hybrid approach that extracts features from fundus images and performs binary and multiclass classifications of DR. Following the pre-processing stage, the images are sent into the transfer learning models. Both models will be used to extract a total of 1000 features. Then, four classifiers will receive the combined 2000 features as input to predict the output classes. SVM can achieve the highest percentage accuracy and other metrics in both binary and multiclass classification using hybrid features extracted from both models.

The authors of this study [62] propose a hybrid model that combines EfficientNet (for local feature extraction of the image) and Swin Transformer (to extract the global features) for the classification of DR staging. An algorithm for automated cropping of the region of interest while the image's center is kept, Gaussian blur to enhance visual clarity and contrast and additional techniques include learning rate scheduling, weight attenuation Dropout, and several data augmentation algorithms. The Model is separated into two branches: EfficientNet is used as the backbone network, while Swin Transformer is used in another branch. The two branches' outputs (extracted local features and global features) are fused through a fully connected layer after being concatenated into the classification layer for the classification of DR stages.

In Aykat and Senan [74], nine different CNN architectures are applied to the OCT retinal dataset to detect 4 retinal disease types and acquire the performance outcomes of each to detect success rates. Then, to improve these models' performance, hyper-parameter tuning is employed. Finally, based on the most effective models of these hyper-parameter tuned architectures, they created a hybrid CNN model (EfficientNetV2S and Xception) named EffXCeptNet. The proposed model was demonstrated through comparisons of the performance results with other models and outperformed them.

The authors of this study Kermany *et al.* [75] suggest a novel DL-based model for accurately predicting various eye conditions from OCT images. It is a two-step procedure that starts with segmentation using a trained U-Net model and ends with disease classification using an ensemble model made up of the Xception and InceptionV3 networks. To improve classification, the self-attention method is used to enhance model ability which makes use of each model's feature maps. The web application is built that permits users to upload their raw OCT scan pictures and obtain disease classification results for OCT images.

The automatic DR detection diagnostic tool proposed in this paper [67] named a hybrid inductive machine learning algorithm (HIMLA) that analyzes colored fundus images and classifies them as either unhealthy (having DR) or healthy (no DR). The input photos undergo pre-processing to remove any unwanted background details and improve contrast and brightness. The processed images are encoded and decoded to segment them during the segmentation stage. Based on significant information from the segmented region and how an image is altered from its original state to one afflicted by a disease, features have been extracted. Classifying retinal images using multiple instance learning (MIL) comes in the last step.

The work presented in Pin *et al.* [70] is based on an ensemble of two CNN transfer learning models. The authors propose an ensemble CNN model to analyze OCT images and classify five retinal conditions. Before feeding OCT images into neural network models, image pre-processing techniques applied, such as background removal, and contrast limited adaptive histogram equalization (CLAHE) are used to improve the quality of the images, image size reduction, and data augmentation to dataset size increasing and overfit preventing. Finally, an ensemble method based on the MobileNetV3Large [76] and ResNet50 [77] fusion probability output is suggested for achieving a more robust OCT image classification performance.

In Ishtiaq *et al.* [59] authors propose a hybrid approach for the identification and classification of DR using an ensemble-optimized CNN. The pre-processing steps include resizing images, augmenting data, median-filter applying, and sharpening images. Three approaches are utilized for feature extraction (LBP, GraphNet124, and ResNet50). Shannon Entropy algorithm is used for the selection of those features. Binary Dragonfly Algorithm [78] and Sine Cosine Algorithm (SCA) [79] are used for fusing and optimizing selected features. Ten ML algorithms are used (five of which are SVM variations and five are KNN variants), and optimized feature vectors are fed to those. Each of these algorithms was assessed using a separate evaluation matrix depicted in Table 3.

The authors in this research [57] present an innovative hybrid framework for Gl detection (Gl or non) combining CNNs (ResNet50, VGG-16) and ML (Random Forest) in an ensemble approach. The methodology consists of

Table 3: Details of chosen studies for eye disease diagnosis

Study (Year)	Dataset	Feature extraction method	Classification model	Type of classification	Accuracy (%)
Abbas *et al.* [45], 2022	• ODIR	• Convolutional layers (CNN)	• CNN	• Multi-class (4) • Two classes	• 96.9 • 100
Khan *et al.* [49], 2021	• ODIR (Kaggle)	• VGG-19	• VGG-19	• Binary classification	• 97.47
Hind Hadi *et al.* [46], 2020	• ODIR	• EfficientNet	• Custom neural network	• Multi-Label multi-class • (8 classes) Classification	• 89
Aljohani and Aburasain [58], 2022	• DRISTHI- GS • ACRIMA-DB	• HOG • GLCM	• LDA • RF • SVM • NB	• Multi-class (4)	• 73.33 • 82.56 • 82.64 • 88.86
Al-Timemy *et al.* [81], 2023	• EyeNet	• CNN	• CNN	• Multi-class (32)	• 95
Nawaz *et al.* [82], 2023	• MuReD	• DenseNet161	• C-Tran	• Multi-class (20)	-
Hossain *et al.* [39], 2023	• IDRiD • ODIR • DRISHTI-GS • Retinal Dataset-GitHub • Messidor • Messidor-2	• ResNet50 • VGG-16 • Xception • EfficientNetB7 • Custom DCNN model	• ResNet50 • VGG-16 • Xception • EfficientNetB7 • Custom DCNN model	• Multi-class (4)	• 96.94
Hossain *et al.* [50], 2024	• DIARETDB0 • DIARETDB1 • Messidor • HEI-MED • ODIR (Kaggle) • Retina	• EfficientNetB0 • VGG16 • ResNet152V2 • GRU+ResNet152V2 • Bi-GRU+ResNet152V2	• EfficientNetB0 • VGG16 • ResNet152V2 • GRU+ResNet152V2 • Bi-GRU+ResNet152V2	• Multi-class (4)	• 98.76 • 98.11 • 97.3 • 98.38 • 98.11
Hossain *et al.* [55], 2020	• STARE	• CNN (AlexNet)	• CNN (AlexNet)	• Multi-class (4)	-
Hossain *et al.* [38], 2020	• Private • HRF • DRIVE • STARE • IDRiD	• DCNN (AlexNet)	• DCNN (AlexNet)	• Binary classification	• 95.77
Vadduri and Kuppusamy [40], 2021	• ORIGA • IDRiD • MESSIDOR • ARIA • STARE	• ResNets • VGGs	• DNN	• Multi-class (4)	• 85.79
Rodríguez *et al.* [83], 2021	• Kaggle	• MobileNetV2	• MobileNetV2 • (Transfer learning)	• Multi-class (5)	• 96.2
Wang *et al.* [47], 2021	• ODIR	• Resnet-34 • EfficientNet • MobileNetV2 • VGG-16	• Resnet-34 • EfficientNet • MobileNetV2 • VGG-16	• Multi-class (8)	• 90.85 • 94.32 • 93.82 • 97.23
Dipu *et al.* [48], 2022	• ODIR	• VGG-19	• VGG-19	• Binary classification	• 94.0 • 88.9 • 86.1 • 86.6 • 98.1 • 90.9 • 86.8
Ali Tabtaba and Ata [43], 2022	• RFMiD	• ML-CNN	• ML-CNN • (sigmoid)	• Multi-class (45)	• 94.3
Vanita Sharon and Saranya [56], 2023	• DRISTHI-GS1 ORIGA	• DNN • (SqueezeNet)	• KNN • SVM • DT • RF • NB • LR	• Binary classification	• 65.3–99

(Contd...)

Table 3: *(Continued)*

Study (Year)	Dataset	Feature extraction method	Classification model	Type of classification	Accuracy (%)
Chea and Nam [41], 2023	• IDRiD • HRF • ODIR • RFMiD	• CNN • Transfer learning (EfficientNet)	• CNN • Transfer learning	• Multi-class (4)	• 84 • 94
Albelaihi and Ibrahim [51], 2021	• Messidor • Messidor-2 • DRISHTI-GS • Kaggle	• CNN	• CNN	• Multi-class (4)	• 81.33
Sarki *et al.* [52], 2019	• Messidor	• AlexNet • VGG-16 • SqueezeNet • Custom CNN	• AlexNet • VGG-16 • SqueezeNet • Custom CNN	• Binary classification	• 93.46 • 91.82 • 94.49 • 98.15
Guo *et al.* [84], 2020	• Kaggle	• CNN • (AlexNet)	• CNN • (AlexNet)	• Binary classification	• 88.13 • 73.75 • 89.38 • 86.88 • 97.50

DRIVE: Digital retinal images for vessel extraction, IDRiD: Indian DR image dataset, HRF: High-resolution fundus, STARE: Structured analysis of the retina, RFMiD: Retinal fundus multi-disease image dataset, ODIR: Ocular disease intelligent recognition, ORIGA: Online retinal fundus image dataset for Gl analysis and research, DIARETDB0: Standard DR database calibration level 0, DIARETDB1: Standard DR database calibration level 1, ARIA: Automatic retinal image analysis, APTOS-2019: Asia Pacific Tele-Ophthalmology Society 2019, LDA: Linear discriminant analysis, CNN: Convolutional neural network, DCNN: Deep: Convolutional neural network, Messidor: Methods to evaluate segmentation and indexing techniques in the field of retinal ophthalmology, Messidor-2: Methods to evaluate segmentation and indexing techniques in the field of retinal ophthalmology-version 2

four main steps including data collection, pre-processing (converting images to gray-scale and texture extraction), training 3 models, and classification. GLCM was used to extract texture features, which were then fed into the Random Forest algorithm. Moreover, retinal grayscale fundus have been fed into the ResNet50 and VGG16 models. By using ensemble modeling (combining 3 models), the Gl detection framework was carefully created.

In Al-Timemy *et al.* [80] authors developed a hybrid DL model to detect keratoconus (KCN) based on corneal maps. To identify KCN-induced lesions associated with a specific corneal map exclusively, they first created seven DL models based on the EfficientNet-b0 architecture, each of which was trained to extract deep features from that map. Next, deep features from various corneal maps are combined to create a hybrid model that combines deep features (1000) extracted from each corneal map based on a support vector machine to generate a concatenated vector with 7000 deep features.

To classify fundus images for 3 disease types, this study [60] uses a hybrid approach that combines ensemble learning and deep-trained feature extraction models. After some pre-processing methods, this study used 8 DL pre-trained models for feature extraction and applied 3 ensemble learning models (Extra Trees, Histogram Gradient Boosting, and Random Forest) for classification. The results showed that the combination of DenseNet for feature extraction and ensemble learning models (classifiers) produced the best results and outperformed other techniques and classifiers. To precisely predict the risk of diabetes at an early stage, a novel Hybrid CNN and Autoencoder model is presented in this paper [69]. The methodology taken by the authors includes data collection, pre-processing (Image Resizing, Label encoding, and Data Augmentation), and Proposing a hybrid model of CNN and autoencoder. The CNN is used to extract spatial features from the retinal images, while Autoencoder is used for unsupervised feature learning and dimensionality reduction, by combining the advantageous characteristics of both a hybrid model is designed.

This study [53] proposed a hybrid DL-metaheuristic model for automated diagnosis of DR. An InceptionV3 DL model, is used for feature extraction from fundus images. Simulated Annealing (SA) is applied in the feature selection process. The proposed model has four steps. (1) Retinal image pre-processing. (2) In the feature extraction step, processed images are fed into a deep CNN network and abstract features are obtained using the transfer learning approach. (3) In the feature selection step, a metaheuristic algorithm (i.e., SA) is applied to reduce the number of features while obtaining the best potential features. (4) A decision tree-based algorithm (XGBoost) which is an ensemble method is used as a classifier.

Another research [36] proposed a hybrid feature extraction method for 4 types of eye disease classification. Pre-processing tasks include ROI Extraction and image resizing to a standardized size. Feature extraction is done by utilizing 3 methods (LBP, GLCM, and TFCM). Then, those features are combined into a single feature vector, representing the texture and statistical features of the eye disorders. Finally, the two classifiers were utilized for classification (kNN and SVM).

A DR diagnosis model is proposed in Ali Tabtaba and Ata [42], in which fundus images are taken from 3 datasets. The pre-processing tasks were done including scaling, cropping, and CLAHE. When image augmentation is performed, high-quality images are provided for additional processing carried out by GANs. The augmented photos are subsequently fed into a hybrid cascaded multi-scale Dilated CNN model (HCMD-CNN), whereby the MobileNet and Residual Attention Network are used to produce a promising result. By suggesting the novel algorithm called MSTGEO, the hyper-parameters in the network were optimally selected.

The authors of this research [63] propose a hybrid DL strategy for the identification and classification of DR, utilizing two VGG network models (VGG16 and VGG19) in conjunction with the deep CNN method. First images were pre-processed and then CNN and two VGG NETWORK models were used for feature extraction and classify the data.

The goal of this work [64] is to apply hybrid strategies based on feature extraction and fusion methods to classify an eye disease dataset. The first approach uses an ANN to classify fundus images utilizing features from the DenseNet121 and MobileNet models independently after using principal component analysis to reduce the high dimensionality and repetitive features. The second approach uses ANN based on fused features from the MobileNet and DenseNet121 models, both before and after feature reduction. The third approach uses ANN to categorize the eye illness dataset using hand-crafted features and fused features from the MobileNet and DenseNet121 models separately. The ANN achieved the highest accuracy by combining handcrafted features with a fused MobileNet.

4. DISCUSSION AND FUTURE WORK

The primary aim of this research is to present a comprehensive review of hybrid models form an important aspect of DL-based system analysis of retinal fundus images. A hybrid model due to the integration of different methods used to diagnose eye diseases.

A summary and discussion of the performances of chosen schemes are provided in this section. ML and DL-based approaches for eye disease identification and classification using retinal images are shown in Table 3, and hybrid-based approaches are depicted in Table 4. The accuracy ranges of ML/DL studies vary between 73% and 98%. The maximum accuracy was 98.76% achieved by Albelaihi and Ibrahim [50]. In which authors proposed a model called "DeepDiabetic" for diagnosing and identifying four different types of diabetic eye diseases, they considered five architectures' performances, and the EfficientNetB0 model did better than the other five which achieved high accuracy. The minimum accuracy was obtained by Nawaldgi and Lalitha [58]. In this proposed work, structural (CDR and RDR) and textural (HOG and GLCM) features are used to develop an automated Gl assessment approach from CFI. Among four Classifiers utilized for the classification task Linear Discriminant Analysis achieved the lowest accuracy (73.33%).

It is a noteworthy achievement to use DL models to analyze retinal fundus images with over 90% accuracy. Among the studies, many methods achieved an accuracy of over 90%, including those reported from [36-39], [41-45], [47], [49,50], [52,53], [57], [59], [63,64], [67-70], [74], [81], [83], and [85]. In contrast, others achieved lower than 90% accuracy, including [40], [46], [51], [58], [60], [66], and [80]. See Table 4, in some of those studies, the authors' focus was on using traditional ML feature extraction methods [58]. However, these might not be sufficient to capture the full complexity of the data when compared with DL-based approaches which are the most commonly used to date. The pre-processing step enhances the effectiveness of DL approaches, especially in analyzing medical images. If the pre-processing stages do not handle imbalanced datasets, image quality enhancement, noise reduction, and other standards used in image pre-processing analysis the model might perform well on the majority of diagnoses but lead to poor feature extraction and fail to detect rare retinal abnormalities. Furthermore, relying on only a single or small dataset might restrict the model's generalization to other datasets or real-world clinical data. Data augmentation can be used to increase the amount of samples. Thus, taking these problems into account will significantly enhance the model's performance.

Table 4 demonstrates hybrid-based methodologies built for eye disease identification and classification using retinal

Table 4: Details of hybrid studies for eye disease diagnosis

Study (Year)	Dataset	Feature extraction method	Classification model	Type of classification	Performance (%)
Doddi *et al.* [36], 2023	• Kaggle (eye disease classification)	• LBP • GLCM • TFCM	• SVM • KNN	• Multi-class (4)	• 99.88 • 99.55
Abbood *et al.* [37], 2024	• Kaggle (EyePACS) • Messidor	• ResNet50	• ResNet50	• Multi-class (5)	• 92
Retina Dataset [66] 2021	• Kaggle	• (CNN) • InceptionV3 • InceptionResNetV2 • DenseNet169	• (RNN) • LSTM	• Multi-class (4)	• 69.50
Khan *et al.* [53] 2023	• Messidor-2	• InceptionV3 • Simulated Annealing	• XGBoost	• Binary classification	• 92.55
Butt *et al.* [62] 2024	• Kaggle • (APTOS-2019)	• EfficientNet • Swin Transformer	• Hybrid model	• Multi-class (5)	• 97
Xu *et al.* [63] 2022	• Kaggle (APTOS-2019)	• CNN • VGG16 • VGG19	• CNN • VGG16 • VGG19	• Multi-class (5)	• 90.60
Londhe [67] 2023	• CHASE	• Optimization of attributes	• Multiple instance learning	• Binary classification	• 96.62
Thanki [57] 2024	• ACRIMA • G1020 • ORIGA • REFUGE	• GLCM	• Random Forest • ResNet50 • VGG16	• Binary classification	• 95.41
Nawaldgi and Lalitha [59] 2023	• Kaggle (EyePACS)	• LBP • GraphNet124 • ResNet50	• SVM • KNN	• Multi-class (5)	• 98.85
Verma *et al.* [61] 2022	• Kaggle (APTOS-2019)	• (Transfer learning) • GoogleNet • ResNet-18	• NB • RF • RBF • SVM	• Binary • Multi-class (3)	• 97.80 • 89.29
Ouda *et al.* [44] 2022	• RFMiD	• EfficientNetB4 • EfficientNetV2S	• Ensemble learning	• Multi-class (4)	• 96.23
Mahmoud *et al.* [68] 2023	• Kaggle	• ResNet-152 • DenseNet-121	• Hybrid model (DRNN)	• Multi-class (27)	• 96.91
Ishtiaq *et al.* [60] 2024	• Kaggle (EyePACS)	• DenseNet201 • InceptionResNetv2 • MobileNetV2 • ReseNet152V2 • NASNetMobile • NASNetLarge • VGG16 • VGG19	• Random Forest • Extra Tree • Histogram Gradient	• Multi-class (3)	• 87.2
Shimpi and Shanmugam [69] 2024	• Private	• CNN • Autoencoder	• CNN • Autoencoder	• Multi-class (5)	• 90.92
Sarode and Desai [70] 2021	• Private	• MobileNetV3Large • ResNet50	• Ensemble model • (CNN)	• Multi-class (5)	• 91.69
Menaouer *et al.* [64] 2023	• Kaggle (eye disease classification)	• DenseNet-121 • MobileNet	• ANN	• Multi-class (4)	• 98.5
He *et al.* [74] 2023	• OCT	• EfficientNetV2S • Xception	• Hybrid model • (EffXCeptNet)	• Multi-class (5)	• 99.90
Syarifah *et al.* [85] 2024	• Public	• InceptionV3 • Xception	• Ensemble model	• Multi-class (4)	• 96.60
Mirjalili [80] 2021	• Private	• (Hybrid model) • EfficientNetB0	• SVM	• Multi-class (5)	• 81.6
Babaqi *et al.* [42] 2024	• IDRiD • DIARETDB1 • Messidor-2	• MobileNet • Residual Attention Network	• MSTGEO-HCMD-CNN	• Multi-class (5)	• 95.79

RFMiD: Retinal fundus multi-disease image dataset, ODIR: Ocular disease intelligent recognition, ORIGA: Online retinal fundus image dataset for GI analysis and research, REFUGE: Retinal fundus GI challenge, APTOS-2019: Asia Pacific Tele-Ophthalmology Society 2019, CNN: Convolutional neural network, Messidor: Methods to evaluate segmentation and indexing techniques in the field of retinal ophthalmology, Messidor-2: Methods to evaluate segmentation and indexing techniques in the field of retinal ophthalmology-version 2, GLCM: Gray-Level Co-Occurrence Matrix

images and OCT images. The obtained accuracy of the proposed hybrid-based models ranges from 69% to 99%. The highest accuracy belongs to Aykat and Senan [74] which is 99.9%, and 69.5% of Londhe [66] is the lowest accuracy of all the chosen classification techniques.

Similarly, in terms of sensitivity (True Positive Rate or Recall) as shown in Fig. 4, the maximum obtained sensitivity scores among all of the chosen researches is 99.9% of Aykat and Senan [74], and the minimum one is 69.5% which belongs to Londhe [66]. A sensitivity score of above 90% was attained in certain works, including [36], [42], [59], [61], [62], [67], [68], [70], [74], [85] and [63], while others reported sensitivity below 90%. In addition, some researches scored sensitivity up to 95%, such as [61] (binary classification), [36], [59], [67], [74], [85] and [42] as shown in Fig. 5.

Regarding acquiring precision scores for chosen researchers, they range from 89.40 % to 99.9%. The study [74] scored the highest precision rate of 99.9% in which they built a hybrid CNN model named EffXCeptNet (combined EfficientNetV2S and Xception), authors applied nine different CNN architectures on OCT retinal dataset to detect 4 retinal disease types. The pre-trained CNN models of GoogleNet and ResNet-18 are used in Butt *et al.* [61] to propose a hybrid approach that extracts features from fundus images and performs binary and multiclass classifications of DR. However, their binary classification achieved a 97.8% precision rate but the multiclass classification achieved the lowest precision rate of 89.40% among all other studies where chosen.

Fig. 6. shows the distribution of papers published from 2019 to 2024. The number of published studies related to retinal diseases was retrieved from PubMed [86]. The number of publications in this field has steadily increased over the years, reflecting the growing interest and advancements in DL applications for fundus image analysis. Notably, the number of studies rose from 13 in 2019 to 73 in 2024, demonstrating an accelerating research focus in this area. The trend suggests that research interest peaked in 2024, likely due to advancements in DL techniques including hybrid approaches and the availability of large-scale datasets.

The most commonly used datasets in retinal disease diagnosis studies are shown in Fig. 7. The findings reveal that Messidor and EyePACS are the most frequently utilized datasets. This is likely because these datasets are widely used in DR diagnosis, which remains a primary focus in retinal disease research due to its high global prevalence. ORIGA is also widely used; it was specifically developed for Gl diagnosis and OD/OC segmentation. The IDRiD dataset was created to assess the severity of DR and DME. It stands out for its high-quality annotations, making it a valuable resource for accurate disease grading and analysis.

Moderately used datasets, such as ODIR and DRIVE, are often chosen for specific diagnostic tasks such as multi-disease classification and Vessel Extraction. On the other hand, RFMiD and HRF are used less frequently in studies, as researchers often prefer datasets that better suit their specific research goals and offer larger sample sizes. These findings highlight a strong preference for large, well-annotated datasets, which are essential for effectively training and

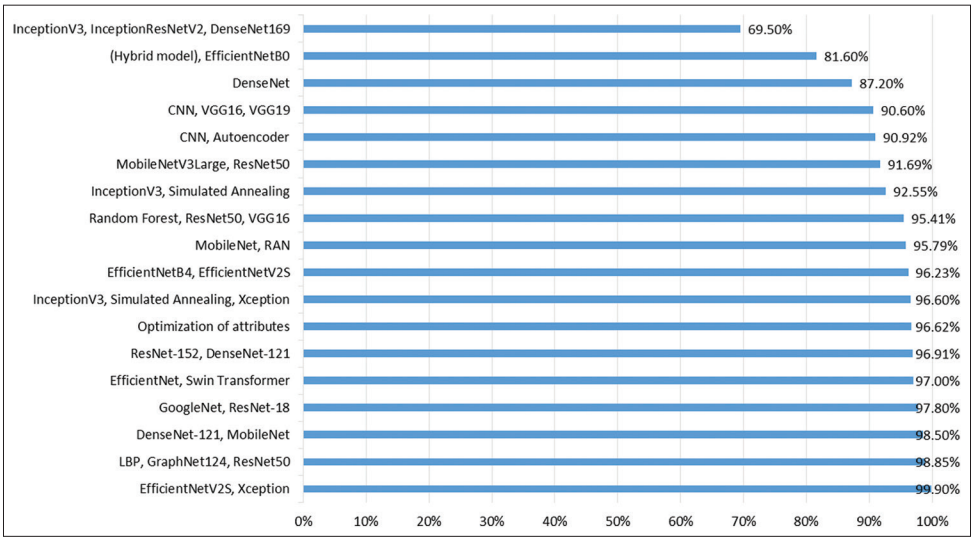


Fig. 4. The effect of hybrid deep learning on performance.

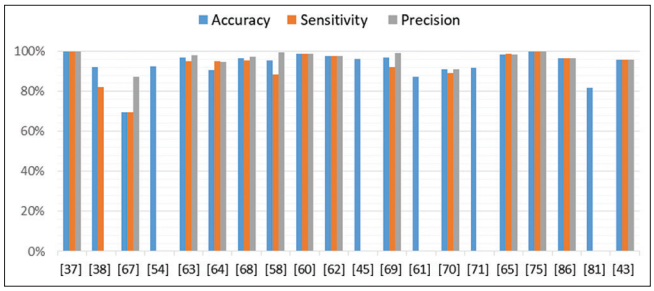


Fig. 5. Accuracy, sensitivity and precision for hybrid deep learning models.

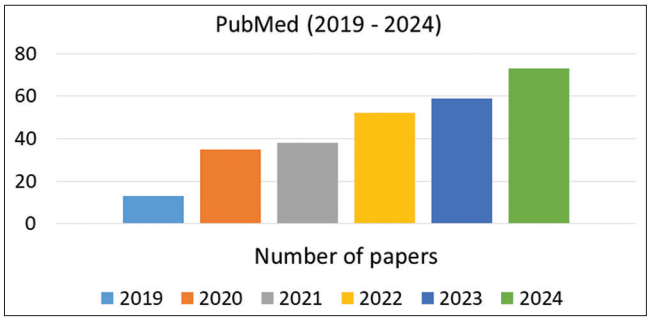


Fig. 6. Number of papers published per year (2019–2024).

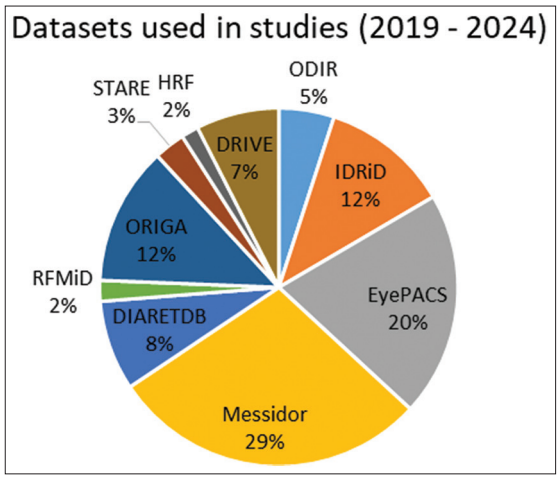


Fig. 7. Percentage of datasets used in papers (2019–2024).

validating DL models in this field. However, the limited use of certain datasets suggests they may be underutilized, which could be addressed in future research to improve dataset diversity and enhance model generalization.

Although the use of DL in fundus image analysis has produced excellent outcomes, it is important to remember that there are still many other areas in which it is still limited. After reviewing various learning-based techniques for the detection and classification of eye diseases, the following outlines their advantages, and limitations of recent studies.

Certain studies primarily focus on performing fundus image detection or classification of one type of eye disease alongside normal (healthy) cases [38], [48], [49], [52], [56], [84], for example, the presence or absence of DR or CAs, that is, binary classification. Exploring the scope to deal with a wider range of conditions could enhance the model's applicability and make it more versatile for real-world clinical use. The process of Disease detection/classification expanded in Ouda *et al.* [43], Wang *et al.* [46], Dipu *et al.* [47], Retina_Dataset [68], Rodríguez *et al.* [82] by considering some other types of Diabetic eye diseases including AMD, CAs, myopia and more. Expanding to contain diverse diseases can elevate the quality of their work much more and make it more impactful and applicable in real-world scenarios. While some datasets have Fundus images with labeling stages of some diseases which can be considered for eye disease detection and severity grading (e.g., mild, moderate, and severe), in this way the studies conducted in Hossain *et al.* [38], Khan *et al.* [49], Khan *et al.* [52], Thanki [56] are significantly enhanced as it enables the system to detect the diseases at earlier stages. The performance of models can be improved by focusing on more practical solutions to address the class imbalance issue present in the suggested datasets, which still needs further attention in several studies [43], [46], [56], [82], [83].

Some studies only utilize a single CNN architecture [49], comparisons with other models or previous researches are not given [51], [55], [84], while it might add significant value to the research. It might be more useful for Nawaz *et al.* [81], Guo *et al.* [83] to compare their approach with a wider range of state-of-art models to highlight the strengths/weaknesses of the proposed approach and could provide more insights for researchers about various models' strengths and effectiveness. While also essential for enriching findings and showing how your work is better and differs from others.

Researchers make use of both ML [36], [58] and DL models for eye disease diagnosis. Each approach offers unique advantages, but DL-based models could be used in Thanki [56], Nawaldgi and Lalitha [58] to automatically extract features and improve classification tasks. In addition, it has shown great success in tasks such as image analysis, particularly in medical image classification.

When datasets have a limited number of samples which is considered low, this might raise the risk of overfitting and class imbalance. Therefore, studies like [42], [59], [64]

have applied strategies to increase fundus images, such as data augmentation to balance the data and DL models to generalize to any new datasets as a large number of samples are required during the training phase. Some studies [46], [55], [83] rely only on a single and small dataset which might increase the restriction of the model's generalization on other datasets or real-world clinical data. On the other hand, studies like [42], [57] evaluated their mode with more than one dataset.

The pre-processing stage is an essential section in research papers because of the details of how raw data in datasets are prepared for feature extraction and training of the models. However, several studies, including [38], [41], [56], [58] do not provide detailed information about pre-processing of dataset images such as how they deal with dataset imbalance, image cleaning, normalization, data augmentation, resizing, and more.

Future work is an important part of research sections that contain valuable information about some works that are missed currently and new ideas that might be considered in future work. Some researchers do not suggest any additional or alternative approaches that might be explored in future research, such as [47], [52].

Some studies [47], [56] seek to classify different types of ocular diseases but they do not provide a brief overview of the various types of ocular diseases that they tried to detect, which might be useful for readers to obtain basic information about them to realize how normal images are different from abnormal ones and which portion of the image is important to determine the exact disease.

To evaluate the effectiveness of the models, it is essential to provide details of performance metrics such as accuracy, precision, recall, and others. The absence of showcasing these metrics, as observed in Vanita Sharon and Saranya [55], leaves the readers without a clear understanding of how well the model performs.

As observed in recent studies, a key trend in this field is the recent advancements of a synthesis of the most recent transfer, ensemble, and DL techniques that are significantly utilized to create an accurate, reliable model for eye disease diagnosis; this could potentially increase accuracy and capacity of predictions. More especially, hybrid DL models offer stronger feature extraction capabilities, making them highly effective for complex classification tasks. Another trend is the transition from single-disease detection to multi-disease classification and severity grading that enables early detection and supports efforts for widespread screening.

Future research should focus on offering an AI-automated tool (Real-time implementation) for diagnosing disorders related to the eyes that experts in medical departments can use. It is desperately needed in clinical centers and serves as an Internet of Medical Things (IoMT) application. This strategy helps with the early referral of emergency patients to specialists and may reduce the time needed for clinical diagnosis.

In addition, expanding the dataset's size or combining several different datasets to increase the number and include a more varied sample of demographics, nationalities, and geographic regions builds a more precise or enhances the model's resilience and generalizability.

5. CONCLUSIONS

This paper proposed a concise review of introducing a range of learning paradigms including DL, hybrid, and ensemble models utilized for the purpose of diagnosis of eye diseases. More specifically, this review focused on one type of ophthalmic imaging modality, which includes an extensive range of disorders relating to the eyes in image format, that is, fundus images. In addition, highlights the advantages of hybrid techniques that integrate different machine learning and DL algorithms with required image processing methods to detect, diagnose, classify, and grading various retinal diseases. A survey of the selected papers was conducted over the past 5 years; here we narrowed down the review of studies conducted in 2019. We described and summarized the methodologies of selected studies in the area of fundus image diagnosis, including data collection, pre-processing approaches, feature extraction, and classification models. According to our analysis of these studies, hybrid techniques have been widely utilized by researchers for the purpose of combining the strengths of different models, enhancing diagnostic accuracy, and improving feature extraction and classification.

In addition, we have noted a number of restrictions related to the investigations. We have discussed their advantages, challenges, emerging trends, and potential opportunities for future research in this field. Furthermore, we highlighted the strengths and weaknesses of the most commonly used datasets in these studies. As we realized AI systems can help physicians by offering an automated tool for the

early identification, classification, and grading of various eye conditions. Given the lack of medical specialists relative to the number of patients, offering a cost-effective automated AI system for diagnosing disorders related to the eyes is desperately needed, which helps support medical professionals and allows patients to begin treatment sooner. We hope that our research will be able to provide an in-depth analysis and cover a thorough and up-to-date summary of Eye disease diagnosis.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the financial support provided by the University of Sulaimani. Special thanks to the Computer Department/College of Science at the University of Sulaimani for providing a suitable environment to fulfill this project.

REFERENCES

[1] WHO. "Blindness and vision impairment." World Health Organization. Available from: https://www.who.int/news-room/fact-sheets/detail/blindness-and-visual-impairment [Last accessed on 2024 Aug 21].

[2] D. Bordoloi, V. Singh, S. Sanober, S. M. Buhari, J. A. Ujjan and R. Boddu, "[Retracted] Deep learning in healthcare system for quality of service," *Journal of Healthcare Engineering,* vol. 2022, no. 1, p. 8169203, 2022.

[3] G. A. Saleh, N. M. Batouty, S. Haggag, A. Elnakib, F. Khalifa, F. Taher, M. A. Mohamed, R. Farag, H. Sandhu, A. Sewelam and A. El-Baz, "The role of medical image modalities and AI in the early detection, diagnosis and grading of retinal diseases: A survey," *Bioengineering,* vol. 9, no. 8, p. 366, 2022.

[4] ophthalmology.med.ubc.ca. "*Color Fundus Photography*," Available from: https://ophthalmology.med.ubc.ca/patient-care/ophthalmic-photography/color-fundus-photography [Last accessed on 2024 Aug 21].

[5] The University of British Columbia. "*Ophthalmology Fundus Image*," Available from: https://ophthalmology.med.ubc.ca/files/2013/11/img-03.jpg [Last accessed on 2025 Feb 11].

[6] M. Mateen, J. Wen, Nasrullah, S. Song and Z. Huang, "Fundus image classification using VGG-19 architecture with PCA and SVD," *Symmetry,* vol. 11, no. 1, p. 1, 2018.

[7] D. Neely. "*Diabetic Retinopathy: Causes, Symptoms, Treatment*," Available from: https://www.orbis.org/en/avoidable-blindness/diabetic-retinopathy?psafe_param=1&gad_source=1&gclid=EAIaIQobChMI4JrP_KCIiAMV0hEGAB1tiA5xEAAYASAAEgLM2fD_BwE [Last accessed on 2024 Aug 22].

[8] "*Glaucoma Overview*," Available from: https://www.nhs.uk/conditions/glaucoma/#:~:text=Glaucoma%20is%20a%20common%20eye,not%20diagnosed%20and%20treated%20early [Last accessed on 2024 Aug 22].

[9] K. Boyd. "*What Are Cataracts?"* American Academy of Ophthalmology. Available from: https://www.aao.org/eye-health/diseases/what-are-cataracts [Last accessed on 2024 Aug 22].

[10] D. Turbert. "*Nearsightedness: What Is Myopia*?" American Academy of Ophthalmology. Available from: https://www.aao.org/eye-health/diseases/myopia-nearsightedness [Last accessed on 2024 Aug 22].

[11] K. Boyd, "*What Is Macular Degeneration*?" American Academy of Ophthalmology. Available from: https://www.aao.org/eye-health/diseases/amd-macular-degeneration [Last accessed on 2024 Aug 22].

[12] P. Porwal, S. Pachade, R. Kamble, M. Kokare, G. Deshmukh, V. Sahasrabuddhe and F. Meriaudeau, "Indian diabetic retinopathy image dataset (IDRiD): A database for diabetic retinopathy screening research," *Data,* vol. 3, no. 3, p. 25, 2018.

[13] A. Budai, R. Bock, A. Maier, J. Hornegger and G. Michelson, "(HRF) Robust vessel segmentation in fundus images," *International Journal of Biomedical Imaging,* vol. 2013, no. 1, p. 154860, 2013.

[14] A. Budai and J. Odstrcilik, *"High-Resolution Fundus (HRF) Image Database*." Pattern Recognition Lab, FAU. Available from: https://www5.cs.fau.de/research/data/fundus-images [Last accessed on 2025 Feb 09].

[15] E. Decencière, X. Zhang, G. Cazuguel, B, Laÿ,…& J. C. Klein, "Feedback on a publicly distributed image database: The messidor database," *Image Analysis and Stereology,* vol. 33, pp. 231-234, 2014.

[16] ADCIS. "*Messidor: A Retinopathy Dataset*." Available from: https://www.adcis.net/en/third-party/messidor [Last accessed on 2025 Feb 09].

[17] J. Staal, M. D. Abramoff, M. Niemeijer, M. A. Viergever and B. V. Ginneken, " Ridge-based vessel segmentation in color images of the retina," *IEEE Transactions on Medical Imaging*, vol. 23, no. 4, pp. 501-509, 2004.

[18] "*STARE: A Retinal Image Dataset*." Available from: http://cecas.clemson.edu/~ahoover/stare [Last accessed on 2025 Feb 09].

[19] S. Pachade, P. Porwal, D. Thulkar, M. Kokare, G. Deshmukh, V. Sahasrabuddhe, L. Giancardo, G. Quellec and F. Mériaudeau, "Retinal fundus multi-disease image dataset (RFMiD): A dataset for multi-disease detection research," *Data,* vol. 6, no. 2, p. 14, 2021.

[20] I. Dataport. "*Retinal Fundus Multi-Disease Image Dataset (RFMID)*," Available from: https://ieee-dataport.org/open-access/retinal-fundus-multi-disease-image-dataset-rfmid [Last accessed on 2025 Feb 09].

[21] "*International Competition on Ocular Disease Intelligent Recognition,*" Available from: https://odir2019.grand-challenge.org/dataset [Last accessed on 2024 Aug 31].

[22] "*Ocular Disease Recognition*," Available from: https://www.kaggle.com/datasets/andrewmvd/ocular-disease-recognition-odir5k [Last accessed on 2024 Aug 31].

[23] "*Messidor2: A Retinopathy Dataset*," Available from: https://www.adcis.net/en/third-party/messidor2 [Last accessed on 2025 Feb 09].

[24] Z. Zhang, F. S. Yin, J. Liu, W. K. Wong, N. M. Tan, B. H. Lee, J. Cheng and T. Y. Wong, "ORIGA(-light): an online retinal fundus image database for glaucoma analysis and research," *Annual International Conference of the IEEE Engineering in Medicine and Biology,* vol. 2010, pp. 3065-3068, 2010.

[25] J. I. Orlando, H. Fu, J. B. Breda, K. van Keer, D. R. Bathula, A. Diaz-Pinto.,& H. Bogunović, REFUGE Challenge: A unified framework for evaluating automated methods for glaucoma assessment from fundus photographs," *Medical Image Analysis,*

vol. 59, p. 101570, 2020.

[26] *R. Challenge. " REFUGE: Retinal Fundus Glaucoma Challenge*," Available from: https://refuge.grand-challenge.org [Last accessed on 2025 Feb 09].

[27] [27] T. Kauppi, V. Kalesnykiene, J.K. Kamarainen, L. Lensu, I. Sorri, H Uusitalo, H. Kalviainen and J. Pietila, "*DIARETDB 0: Evaluation Database and Methodology for Diabetic Retinopathy Algorithms*," 2007.

[28] "*DIARETDB0: A Database for Diabetic Retinopathy*," LUT University. Available from: https://www.it.lut.fi/project/imageret/diaretdb0 [Last accessed on 2025 Feb 09].

[29] [29] T. Kauppi, V. Kalesnykiene, J.K. Kamarainen, L. Lensu, I. Sorri, A. Raninen, R. Voutilainen, H Uusitalo, H. Kalviainen and J. Pietila, "The DIARETDB1 Diabetic Retinopathy Database and Evaluation Protocol. In" "*British Machine Vision Conference*," 2007.

[30] J. Sivaswamy, S. R. Krishnadas, G. D. Joshi, M. Jain and A. U. S. Tabish, "Drishti-GS: Retinal Image Dataset For Optic Nerve Head (ONH) Segmentation. In: "*2014 IEEE 11th International Symposium on Biomedical Imaging (ISBI)*," pp. 53-56, 2014.

[31] "*Drishti-GS: Glaucoma Dataset*." CVIT, IIIT Hyderabad. Available from: http://cvit.iiit.ac.in/projects/mip/drishti-gs/mip-dataset2/Home.php [Last accessed on 2025 Feb 10].

[32] J. Sivaswamy, A. Chakravarty, G. D. Joshi and T. A. Syed, *"(Drishti-GS1)A Comprehensive Retinal Image Dataset for the Assessment of Glaucoma from the Optic Nerve Head Analysis*," [Working Paper], 2015.

[33] J. Emma Dugas, W. Cukierski, "*Diabetic Retinopathy Detection*." Available from: https://www.kaggle.com/c/diabetic-retinopathy-detection/data [Last accessed on 2024 Sep 02].

[34] Kaggle, "*APTOS-2019 Dataset*," Available from: https://www.kaggle.com/datasets/mariaherrerot/aptos2019 [Last accessed on 2024 Sep 02].

[35] G. V. Doddi. Available from: https://www.kaggle.com/datasets/gunavenkatdoddi/eye-diseases-classification/data [Last accessed on 2024 Feb 09].

[36] S. Abd Kadum, F. H. Najjar, H. M. Al-Jawahry, and F. Mohamed, Eye Diseases Classification Based on Hybrid Feature Extraction Methods. In: "*2023 6th International Conference On Engineering Technology and Its Applications (IICETA)*,"IEEE, United States, pp. 402-407, 2023.

[37] [37] S. H. Abbood, H. N. A. Hamed, M. S. M. Rahim, A. Rehman, T. Saba, and S. A. Bahaj, "Hybrid Retinal Image Enhancement Algorithm for Diabetic Retinopathy Diagnostic Using Deep Learning Model," *IEEE Access*, vol. 10, pp. 73079-73086, 2022.

[38] M. R. Hossain, S. Afroze, N. Siddique and M. M. Hoque, "Automatic Detection of Eye Cataract Using Deep Convolution Neural Networks (DCNNs). In: *2020 IEEE Region 10 Symposium (TENSYMP)*, IEEE, United States, pp. 1333-1338, 2020.

[39] M. Vadduri and P. Kuppusamy, "Enhancing ocular healthcare: Deep learning-based multi-class diabetic eye disease segmentation and classification," *IEEE Access,* vol. 99, p. 1, 2023.

[40] N. Chea and Y. Nam, "Classification of fundus images based on deep learning for detecting eye diseases," *Computers, Materials and Continua,* vol. 67, no. 1, pp. 411-426, 2021.

[41] T. Babaqi, M. Jaradat, A. E. Yildirim, S. H. Al-Nimer and D. Won, "Eye Disease Classification Using Deep Learning Techniques," *arXiv preprint arXiv:2307.10501,* 2023.

[42] A. Ali Tabtaba and O. Ata, "Diabetic retinopathy detection using developed hybrid cascaded multi-scale DCNN with hybrid heuristic strategy," *Biomedical Signal Processing and Control,* vol. 89, p. 105718, 2024.

[43] O. Ouda, E. AbdelMaksoud, A. Abd El-Aziz and M. Elmogy, "Multiple ocular disease diagnosis using fundus images based on multi-label deep learning classification," *Electronics,* vol. 11, no. 13, p. 1966, 2022.

[44] R. Abbas, S. O. Gilani and A. Waris, "Ensemble Based Multi-Retinal Disease Classification and Application with Rfmid Dataset Using Deep Learning," *SSRN Journal [Preprint]*, 2022.

[45] A. Hind Hadi, A. S. Ali Yakoob and H. Enass, "Classifying three stages of cataract disease using CNN," *Journal of University of Babylon,* vol. 30, no. 3, pp. 150-167, 2022.

[46] J. Wang, L. Yang, Z. Huo, W. He and J. Luo, "Multi-label classification of fundus images with efficientNet," *IEEE Access,* vol. 8, pp. 212499-212508, 2020.

[47] N. M. Dipu, S. A. Shohan and K. Salam, "Ocular disease detection using advanced neural network based classification algorithms," *Asian Journal For Convergence In Technology (AJCT)*, vol. 7, no. 2, pp. 91-99, 2021.

[48] M. S. Khan., N. Tafshir, K. N. Alam, A. R. Dhruba, M. M. Khan, A. A. Albraikan, F. A. Almalki, "Deep learning for ocular disease recognition: An inner-class balance," *Computational Intelligence and Neuroscience,* vol. 2022, p. 5007111, 2022.

[49] M. S. M. Khan, M. Ahmed, R. Z. Rasel and M. M. Khan, Cataract Detection Using Convolutional Neural Network with VGG-19 Model. In: "*2021 IEEE World AI IoT Congress (AIIoT)*," pp. 209-212, 2021.

[50] A. Albelaihi and D. M. Ibrahim, "Deep diabetic: An identification system of diabetic eye diseases using deep neural networks," *IEEE Access*, vol. 12, pp. 10769-10789, 2024.

[51] R. Sarki, K. Ahmed, H. Wang, Y. Zhang and K. Wang, "Convolutional neural network for multi-class classification of diabetic eye disease," *EAI Endorsed Transactions on Scalable Information Systems,* vol. 9, no. 4, pp. 1-11, 2021.

[52] S. H. Khan, Z. Abbas and S. D. Rizvi, Classification of Diabetic Retinopathy Images Based on Customised CNN Architecture. In: "*2019 Amity International Conference on Artificial Intelligence (AICAI)*," IEEE, United States, pp. 244-248, 2020.

[53] Ö. F. Gürcan, U. Atıcı and Ö. F. Beyca, "A hybrid deep learning-metaheuristic model for diagnosis of diabetic retinopathy," *Gazi University Journal of Science,* vol. 36, no. 2, pp. 693-703, 2023.

[54] A. D. Hoover, V. Kouznetsova and M. Goldbaum, "(STARE)Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response," *IEEE Transactions on Medical Imaging,* vol. 19, no. 3, pp. 203-210, 2000.

[55] A. Vanita Sharon and G. Saranya, Classification of Multi-retinal Disease Based on Retinal Fundus Image Using Convolutional Neural Network. In: A. M. Iliyasu, R. Bestak, and F. Shi Eds. "*New Trends in Computational Vision and Bio-inspired Computing: Selected works presented at the ICCVBIC 2018, Coimbatore, India*," Springer International Publishing, Cham, pp. 1009-1016, 2020.

[56] R. Thanki, "A deep neural network and machine learning approach for retinal fundus image classification," *Healthcare Analytics,* vol. 3, p. 100140, 2023.

[57] A. Aljohani and R. Y. Aburasain, "A hybrid framework for glaucoma detection through federated machine learning and deep learning models," *BMC Medical Informatics and Decision Making,* vol. 24, no. 1, p. 115, 2024.

[58] S. Nawaldgi and Y. S. Lalitha, "Automated glaucoma assessment

from color fundus images using structural and texture features," *Biomedical Signal Processing and Control,* vol. 77, p. 103875, 2022.

[59] U. Ishtiaq, E. Abdullah and Z. Ishtiaque, "A Hybrid technique for diabetic retinopathy detection based on ensemble-optimized CNN and texture features," *Diagnostics (Basel),* vol. 13, no. 10, p. 1816.

[60] J. Verma, I. Kansal, R. Popli, V. Khullar, D. Singh, M. Snehi and R. Kumar. A Hybrid images deep trained feature extraction and ensemble learning models for classification of multi disease in fundus images. In: M. Särestöniemi, *et al.*, Eds., "*Digital Health and Wireless Solutions*," Springer Nature, Switzerland, pp. 203-221, 2024.

[61] M. M. Butt, D. A. Iskandar, S. E. Abdelhamid, G. Latif and R. Alghazo, "Diabetic retinopathy detection from fundus images of the eye using hybrid deep learning features," *Diagnostics,* vol. 12, no. 7, p. 1607, 2022.

[62] H. Xu, X. Shao, D. Fang and F. Huang, "A hybrid neural network approach for classifying diabetic retinopathy subtypes," *Frontiers in Medicine,* vol. 10, p. 1293019, 2024.

[63] B. Menaouer, Z. Dermane, N. El Houda Kebir and N. Matta, "Diabetic retinopathy classification using hybrid deep learning approach," *SN Computer Science,* vol. 3, no. 5, p. 357, 2022.

[64] A. Shamsan, E. M. Senan and H. S. A. Shatnawi, "Automatic classification of colour fundus images for prediction eye disease types based on hybrid features," *Diagnostics (Basel),* vol. 13, no. 10, p. 1758, 2023.

[65] "*Retina_Dataset*." Available from: https://github.com/yiweichen04/retina_dataset/tree/master [Last accessed on 2024 Sep 03].

[66] M. Londhe, "*Classification of Eye Diseases Using Hybrid CNN-RNN Models*," National College of Ireland, Dublin, 2021.

[67] M. H. Mahmoud, S. Alamery, H. Fouad, A. Altinawi and A. E. Youssef, "An automatic detection system of diabetic retinopathy using a hybrid inductive machine learning algorithm," *Personal and Ubiquitous Computing,* vol. 27, no. 3, pp. 751-765, 2023.

[68] J. K. Shimpi and P. Shanmugam, "A Hybrid diabetic retinopathy neural network model for early diabetic retinopathy detection and classification of fundus images," *Traitement du Signal,* vol. 40, no. 6, pp. 2711-2722, 2023.

[69] H. J. Sarode and D. Desai, "Customized mechanism for diabetic risk prediction: A Hybrid CNN–autoencoder approach with emphasis on retinal imaging in the elderly," *Journal of Electrical Systems,* vol. 20, no. 1s, pp. 190-199, 2024.

[70] K. Pin, Y. Nam, S. Ha and J. W. Han, "Deep Learning Based on Ensemble to Diagnose of Retinal Disease using Optical Coherence Tomography. In: "*2021 International Conference on Computational Science and Computational Intelligence (CSCI)*," pp. 661-664, 2021.

[71] I. Bouchrika, "*How to Write Research Methodology in 2024: Overview, Tips, and Techniques*," Available from: https://research.com/research/how-to-write-research-methodology [Last accessed on 2024 Sep 10].

[72] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke and A. Rabinovich, Going Deeper with Convolutions. In: "*2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*," pp. 1-9, 2015.

[73] K. He, X. Zhang, S. Ren, and J. Sun, Deep Residual Learning for Image Recognition, In: "*2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*," pp. 770-778, 2016.

[74] S. Aykat and S. Senan, "Advanced detection of retinal diseases via novel hybrid deep learning approach," *Traitement du Signal,* vol. 40, no. 6, pp. 2367-2382, 2023.

[75] D. S. Kermany, M. Goldbaum, W. Cai, C. C. S. Valentim, H. Liang, S. L. Baxter, A. McKeown, G. Yang, X. Wu, F. Yan,…& K. Zhang, "Identifying Medical Diagnoses and Treatable Diseases by Image-Based Deep Learning," *Cell,* vol. 172, no. 5, pp. 1122-1131.e9, 2018.

[76] D. Haase and M. Amthor, Rethinking Depthwise Separable Convolutions: How Intra-Kernel Correlations Lead to Improved MobileNets. In: "*2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*," pp. 14588-14597, 2020.

[77] [77] A. Howard, M. Sandler, B. Chen, W. Wang, L. C. Chen, M. Tan, G. Chu, V. Vasudevan, Y. Zhu, R. Pang, H. Adam and Q. Le, *et al.* Searching for MobileNetV3. In: "*2019 IEEE/CVF International Conference on Computer Vision (ICCV)*," pp. 1314-1324, 2019.

[78] M. M. Mafarja, D. Eleyan, I. Jaber, A. Hammouri and S. Mirjalili, Binary Dragonfly Algorithm for Feature Selection. In: "*2017 International Conference on New Trends in Computing Sciences (ICTCS)*," pp. 12-17, 2017.

[79] S. Mirjalili, "SCA: A Sine Cosine Algorithm for solving optimization problems," *Knowledge-Based Systems,* vol. 96, pp. 120-133, 2016.

[80] A. H. Al-Timemy, Z. M. Mosa, Z. Alyasseri, A. Lavric, M. M. Lui, R. M. Hazarbassanov, S. Yousefi, "A Hybrid deep learning construct for detecting keratoconus from corneal maps," *Translational Vision Science and Technology,* vol. 10, no. 14, p. 16, 2021.

[81] A. Nawaz, T. Ali, G. Mustafa, M. Babar and B. Qureshi, "Multi-class retinal diseases detection using deep CNN with minimal memory consumption," *IEEE Access,* vol. 11, pp. 56170-56180, 2023.

[82] M. A. Rodríguez, H. AlMarzouqi and P. Liatsis, "Multi-label retinal disease classification using transformers," *IEEE Journal of Biomedical and Health Informatics,* vol. 27, no. 6, pp. 2739-2750, 2023.

[83] C. Guo, M. Yu and J. Li, "Prediction of different eye diseases based on fundus photography via deep transfer learning," *Journal of Clinical Medicine,* vol. 10, no. 23, p. 5481, 2021.

[84] M. A. Syarifah, A. Bustamam and P. P. Tampubolon, Cataract classification based on fundus image using an optimized convolution neural network with lookahead optimizer. In: "*AIP Conference Proceedings*,"vol. 2296, no. 1, AIP Publishing, United States, 2020.

[85] G. Naik, N. Narvekar, D. Agarwal, N. Nandanwar and H. Pande, "Eye disease prediction using ensemble learning and attention on OCT scans. In: "*Future of Information and Communication Conference*," Springer, Germany, pp. 21-36, 2024.

[86] National Library of Medicine "*PubMed Database*." Available from: https://pubmed.ncbi.nlm.nih.gov [Last accessed on 2025 Feb 14].

ORIGINAL RESEARCH ARTICLE UHD JOURNAL OF SCIENCE AND TECHNOLOGY

Optimization of Lattice-Based Cryptographic Key Generation using Genetic Algorithms for Post-Quantum Security

Tara Nawzad Ahmad Al Attar, Rebaz Nawzad Mohammed

Department of Computer Science, College of Science, University of Sulaimani, Sulaymaniyah, Iraq.

ABSTRACT

The progress of quantum computing has posed serious threats to classical cryptographic systems, necessitating much research into developing post-quantum cryptography (PQC). Of the schemes available in PQC, the strongest candidates appear to be lattice-based cryptography (LBC), which encompasses an ample security basis and good computation efficiency. However, practically implementing LBC is faced with key-generation and optimization difficulties, mainly because of its enormous key sizes and computational overhead. The research proposes a novel concept whereby genetic algorithms (GAs) are blended with LBC to increase the merits of key generation while guaranteeing security. Through the evolutionary capacity of GAs, the proposed method optimizes lattice-based keys through selection, crossover, and mutation to ensure high entropy and computationally feasible with experimental results indicating that the GA-based method can cut down memory requirements and computational complexity, making it favorable for resource-constrained environments such as the Internet of Things and embedded systems. The method thus suggested accelerates encryption speed and simultaneously strengthens the security of the optimized key structures. This study emphasizes evolutionary algorithms' potential to facilitate PQC advancement and provides a scalable and efficient framework for cryptographic systems.

Index Terms: Cryptographic Security, Genetic Algorithm, Key Optimization, Lattice-Based Cryptography, Post-Quantum Cryptography

1. INTRODUCTION

The launch of quantum computing has considerably challenged the standard cryptographic setups [1]. Existing classical cryptographic systems, such as RSA and Elliptic Curve Cryptography (ECC) [2], bank on mathematical complexities that are computationally very hard for standard computers but can be quite simply solved by quantum algorithms, such as Shor's algorithm [3]. This weakness has sped up the search for post-quantum computing (PQC) solutions and eventually led to the design of cryptosystems that are amenable to attack by quantum computers [4]. Lattice-based-cryptography (LBC) [5], with its robust security guarantees and great ease of computation has since been in the limelight. In this spirit, the National Institute of Standards and Technology (NIST) [6] has broadened its interests in forming global efforts to standardize PQC algorithms: Some, like lattices, are under scrutiny to become future standards for cryptographic schemes.

Lattice-based cryptography (LBC) is based on the extreme hardness of two computationally challenging lattice problems called the Shortest Vector Problem (SVP) and the Learning

Access this article online

DOI:10.21928/uhdjst.v9n1y2025.pp93-105

E-ISSN: 2521-4217
P-ISSN: 2521-4209

Corresponding author's e-mail: tara.ahmad@univsul.edu.iq

Received: 02-03-2025 Accepted: 27-03-2025 Published: 20-04-2025

with Errors (LWE) [7]. This kind of cryptography is considered a potent building block for digital signatures, key agreements, and encryption security for the quantum age [8]. However, the practical deployment of such systems faces challenges in key generation and optimization. Key sizes in large key lattice systems may yield troublesome memory sizing. At the same time, computational cost may restrict their applicability in low-resource environments such as the Internet of Things (IoT) and embedded systems [9].

LBC derives its security from the computational intractability of certain fundamental lattice problems, primarily the SVP and the LWE problem [7]. The SVP involves finding the shortest non-zero vector in a high-dimensional lattice, which is known to be NP-hard under specific norms [10]. The LWE problem, on the other hand, is based on solving linear equations with small random errors, making it resistant to both classical and quantum attacks. These problems serve as the basis for constructing secure encryption schemes, digital signatures, and key exchange protocols [11]. Unlike factorization-based cryptosystems, which are vulnerable to Shor's algorithm, the hardness of SVP and LWE remains intact even in the presence of quantum computers [12]. As a result, LBC has emerged as a leading candidate for PQC and is actively being standardized by NIST for future cryptographic applications.

To tackle these problems, genetic algorithms (GAs) are used in this study to address the optimization incorporating lattice-based cryptographic keys. GAs [13], which were taken from biology around 1970, are an efficient and effective technique for solving optimization problems in complex algebraic and other structures. GAs have, in so many words, implemented all the desirable features that were essential to hyperplanes' construction [14], such as characters, crossover/mutation, strategies, concealing, and non-linearities. Through the practice of GA, it is possible to refine lattice-based cryptographic keys by adopting an evolutionary multi-objective technique selection, crossover, and mutation. The GA and LBC concepts meshing enhance the key generation process by pruning the infeasible keyword combinations that do not guarantee high security.

Major Contributions:

- The paper introduces GAs in LBC modification, stating clear processes for obtaining security
- Our proposed change advances the efficiency of key generation, thus reducing resource waste for computation while guaranteeing security
- We mean improving key optimization, memory efficiency, and encryption speed and illustrate how the proposed method performs better for this particular purpose
- This article examines this technique within real-world cryptography systems resistant to quantum computer attacks, focusing on particularly resource-limited environments.

The remaining sections of this paper are as follows: Section 2 provides a comprehensive survey of PQC and lattice-based schemes and explanatory notes on optimization. Section 3 shows how GAs are combined with lattice-based key generation and explains the research strategy used. Section 4 describes an experimental study on the improvement method, including the performance measurement and the results obtained. Finally, Section 5 provides closure to the text and delves into the prospective directions of the study's development.

2. LITERATURE REVIEW

A serious threat to traditional cryptographic schemes by quantum computers has spurred the development of PQC – an arms race for creating cryptographic algorithms capable of withstanding quantum-level attacks. The quest for more robust encryption methods in the future of quantum computing has led the NIST to assume a leading role in standardizing such algorithms. It then seeks an extensive account of recent pasta and relevant methodologies of PQC in LBC, its applications, improved performance, and enhanced security against threats from quantum computers.

Quantum computers use quantum mechanical principles to perform tasks at exponentially faster speeds than classical ones. Notably, integer factorization, which is an example of the task, could be broken through the use of Shor's algorithm, thus making the most used cryptographic systems, such as RSA and ECC "dead." Although Shor's algorithm has only been deduced by theory, it is powerful or efficient enough to continue with any practical quantum computer for big-sized everyday applications. Thus, the emergence of PQC would keep it secure even against the quantum era. Standardization about resistance to quantum algorithms is yet crucial, and even candidates have been evaluated at NIST primarily based on complex mathematical problems, e.g., LBC [15].

The rapid advancements in quantum computing pose significant threats to traditional cryptographic systems,

necessitating a transition toward PQC. UAVs, which rely on open wireless channels and have constrained computational resources, are particularly vulnerable to quantum-based attacks, making PQC integration crucial for securing their communications [16]. The NIST PQC standardization process has driven research toward hardware implementations, leading to the development of PQC hardware circuits and system design techniques to enhance security in various applications [17]. However, the migration of IP networks to PQC remains a complex challenge due to inconsistencies in migration steps, terminology, and security concerns, with present implementations being largely experimental [18]. A structured transition plan to PQC is essential, incorporating identification, protection, detection, and response phases to ensure a quantum-resilient cybersecurity infrastructure [19]. While hybrid cryptographic systems combining traditional and quantum-resistant techniques are being explored, significant challenges remain in practical deployment and regulatory alignment.

One of the candidates in the post-quantum computing (PQC) contest, LBC, has become a beautiful choice. LBC entails the solving of lattice-based problems such as the closest vector problem, the flexible LWE, and the ring LWE issues that are known to be classically and quantumly hard to solve [20]. LBC-based protocols are available for encryption, key exchange, digital signature, and even homomorphic encryption, and they can be used to secure various kinds of data and communication channels. Even though LBC is safe in practice, some real-world problems arise while trying to implement this scheme, where primary key and memory management structures may be inefficient, making it not convenient for use in areas where resources, such as embedded systems and IoT equipment, are few.

The global spread of IoT devices and the subsequent advancement in capabilities of mobile networks to 5G has necessitated the search and development of cryptosystems that will resist attacks facilitated by a quantum computer. IoT devices are of limited computational power, which requires lightweight cryptographic solutions that provide resistance to quantum attacks and are not stretched when it comes to performance [21]. Such a lightweight LBC instance is beneficial in such a scenario as it provides almost the same level of security as the ordinary LBC with approaches that decrease both memory usage and the level of computing resources in general [22]. In particular, a recent study provided efficient symmetric and asymmetric quantum-safe algorithms for an IoT environment, showing significant performance and memory cost improvements over the existing NIST-approved algorithms [21].

With its abundance of highly complex operations, the prospect of LBC has put a lot of pressure on its developers to improve performance [23]. Various forms of vectorization have been instrumental in enhancing lattice-based schemes, namely key exchange protocols. The section has improved and reduced time and energy consumption [24]. Finally, attention was focused on the problems associated with enlarging keys and using lattice-based systems. An example of such systems is NTRU systems, which can provide faster encryption and key generation than RSA but also suffer from the problem of the large sizes of the keys, which renders them unsuitable for many users [25].

LBC has the potential to address challenging security issues, but great care is still required to protect all such systems from the effects of quantum computers. There continue to be hazards associated with lattice structures and their operations, irrespective of whether they are hardware or software components. For that reason, research in this area is necessary to improve the security for LBC and come up with workable measures in the wake of new threats, such as those that may rely on stealth, as seen in side channel methodologies or physical access [26]. Moreover, the transformation of NIST's PQC faceless approach has to consider the security necessity along with the practicality of memory consumed, scheduler cycles utilized, and power consumed by the devices to ensure that the said tablets satisfy the limits of different hardware equipment [27].

Cryptographic systems that are lattice-based have arisen as the answer to the problem of post-quantum vulnerability. This security concern is determined by the specific algorithms encountered while encrypting information. Although the security model of such systems is rather strong, certain complexity is observed in their practical implementation in environments where resources are limited, such as the IOTs or Edge processing. Optimizing LBC production or incorporating other LBC optimal techniques, such as vectorization, has dramatically increased the ability of LBC to adapt to different appliances. Moreover, it may be expedient to evaluate the security, performance, and other parameters in scale as pioneers of this technology accept other changes that are inevitable toward quantum computers. NIST's attempts to invent Projected Quantum Computation Standalone systems for the future are significant for the following safety standards needed. A comparison of the most relevant related

works is provided in Table 1 to highlight the differences and advantages of the proposed method.

3. MATERIALS AND METHODS

3.1. PQC and LBC

PQC [28] is a field of cryptographic research that aims to develop secure cryptographic algorithms resilient to attacks by quantum computers. Classical cryptographic schemes such as RSA and ECC rely on problems, such as integer factorization and discrete logarithms, which are efficiently solvable using Shor's algorithm on a sufficiently powerful quantum computer. PQC schemes are designed to counteract this vulnerability based on problems that remain computationally intractable even in the presence of quantum adversaries.

LBC has emerged as one of the most promising candidates for post-quantum secure cryptographic protocols. It relies on the computational hardness of lattice problems, such as the SVP and the LWE problem. Given a basis $B = \{b_1|, b_2|...,b_n|\}$ of an n-dimensional lattice Λ, the SVP requires finding the shortest non-zero vector $v \in \Lambda$ which is computationally difficult even for quantum computers. The LWE problem, introduced by Regev, is formulated as follows: given a set of linear equations with small noise, recovering the original secret vector s is assumed to be hard. Mathematically, LWE can be defined as Eq(1).

$$A \cdot s + e = b \bmod q \tag{1}$$

Where $A \in \mathbb{Z}_q^{m \times n}$ is a randomly chosen matrix, $s \in \mathbb{Z}_q^n$ is the secret vector, and $e \in \mathbb{Z}_q^m$ is an error term sampled from a small noise distribution. Recovering s given (A,b) is believed to be computationally infeasible for large enough parameters. LBC provides efficient and scalable constructions for encryption, digital signatures, and key exchange protocols.

3.2. GA

GA [13] is an evolutionary optimization technique inspired by natural selection and genetics principles. It is widely employed to solve complex optimization problems where traditional mathematical or heuristic approaches may be inefficient. GA operates on a population of candidate solutions, evolving them iteratively through selection, crossover, and mutation operations to enhance solution quality. The fitness function, denoted as f(x), evaluates the suitability of each candidate solution x within the search space.

The algorithm begins with an initial population $P = \{x_1, x_2,\ldots,x_n\}$, where each x_i is a chromosome encoding a potential solution. Selection is performed to retain high-fitness individuals, commonly using roulette wheel selection, tournament selection, or rank-based selection. In roulette wheel selection, the probability of selecting an individual x_i is proportional to its fitness:

$$P(x_i) = \frac{f(x_i)}{\sum_{j=1}^{n} f(x_i)} \tag{2}$$

Crossover, or recombination, involves exchanging genetic material between selected parents to create offspring. A common approach is single-point crossover, where a random crossover point c is chosen, and genes are swapped between two parents to produce new chromosomes:

$$Child_1 = (Parent_1[: c], Parent_2 [c:]) \tag{3}$$

$$Child_2 = (Parent_2[: c], Parent_1 [c:]) \tag{4}$$

Mutation introduces small random alterations to maintain genetic diversity and prevent pre-mature convergence. If x_i is a binary-encoded chromosome, a bit-flip mutation at position j can be expressed as:

$$x_{ij} = 1\text{-}x_{ij} \tag{5}$$

Where x_{ij} is the j^{th} gene of chromosome x_i. The algorithm iterates through these operations until a termination criterion is met, such as achieving a pre-defined fitness threshold or reaching a maximum number of generations.

3.3. Proposed Method

This study presents a hybrid cryptographic framework hybrids GA and lattice-based encryption to enhance security and adaptability. The method consists of four main stages: key optimization, encryption, mutation, and decryption. Key optimization is performed using GA to generate an invertible encryption key matrix K over Z_{256}. The fitness function ensures the invertibility of K by verifying that its determinant satisfies $gcd(det(K),256) = 1$ and that its modular inverse exists in modulo 256. The fitness function is defined as:

$$F(K) = \begin{cases} \sum K, if\ \gcd\left(\det(k), 256\right) = 1 \\ -\infty, otherwise \end{cases} \tag{6}$$

Where $\sum K$ represents the sum of all elements in the key matrix. The GA iteratively refines K using selection,

crossover, and mutation to obtain an optimal encryption key. Encryption transforms the plaintext message P into its numerical representation and computes the ciphertext C using matrix multiplication in modular arithmetic:

$$C = P \cdot K \bmod 256 \quad (7)$$

Where padding is applied to align with encryption parameters. To enhance security, a GA-based mutation process introduces controlled randomness by flipping bits in C, further obfuscating the encrypted text and increasing resistance against attacks. Decryption is performed using the modular inverse of the key matrix, ensuring correct retrieval of the original message:

$$P' = C \cdot K^{-1} \bmod 256 \quad (8)$$

Where P' ideally reconstructs the original plaintext. The invertibility of K ensures accurate recovery of P without information loss. Integrating GA-based optimization with lattice cryptography provides a robust mechanism resistant to conventional attacks. The introduced noise through mutation adds a layer of obfuscation. The proposed scheme offers a potential post-quantum alternative, leveraging the hardness of modular matrix inversion while maintaining computational efficiency.

To address the potential security vulnerabilities associated with using GAs for cryptographic key generation, we acknowledge that while GAs offer efficiency improvements, they may introduce risks. Specifically, the optimization process might lead to patterns or predictable sequences in the generated keys, which could be exploited by attackers. This concern arises because GAs are inherently stochastic, and if not carefully designed, they may converge to weak or vulnerable keys, thereby reducing the security of the encryption system. In addition, if the search space is not sufficiently large or diverse, or if the genetic operators (such as selection, crossover, and mutation) are not effectively applied, the randomness required for secure cryptographic keys could be compromised, making them more susceptible to cryptanalysis techniques such as brute-force or statistical attacks.

To mitigate these risks, we have incorporated mechanisms within our GA-based key generation process to preserve randomness and avoid pre-mature convergence. This includes ensuring that the search space remains large and diverse, and that proper randomization techniques are used. We also recommend conducting further security evaluations, such as testing against known cryptographic attacks (e.g., differential and linear cryptanalysis) and assessing the strength of GA-generated keys in realistic environments. While GAs can enhance the efficiency of cryptographic key generation, it is essential to carefully design and test the approach to prevent the introduction of vulnerabilities and ensure its applicability in secure cryptographic systems. The flowchart of the complete process is illustrated in Figure 1, providing a visual representation of the step-by-step workflow.

3.4. Computational Complexity Analysis

The computational complexity of the proposed method, "Optimization of Lattice-Based Cryptographic Key Generation Using Genetic Algorithms for Post-Quantum Security," is evaluated by considering the key stages: key optimization using GAs, encryption, mutation, and decryption. In the key optimization stage, the GA iterates over populations of candidate solutions. Selection evaluates each matrix based on the fitness function, which requires computing the determinant and verifying the invertibility condition. The time complexity of this operation is $O(n^3)$. Crossover combines matrices with complexity $O(n^2)$, and mutation flips bits in the matrix, also requiring $O(n^2)$. Evaluating the fitness function involves determinant and modular inverse calculations, leading to a total complexity of $O(n^3 + n^2)$ per generation.

For encryption, the process involves matrix multiplication between the plaintext and the key matrix, with a complexity of $O(n^2)$. Similarly, the mutation step in the ciphertext involves flipping bits, which also takes $O(n^2)$. Decryption requires matrix multiplication with the inverse of the key matrix, along with modular inverse computation, resulting in a complexity of $O(n^3 + n^2)$.

The overall time complexity of the method is dominated by the GA key optimization and can be expressed as $O(g \cdot p \cdot (n^3 + n^2))$, where g is the number of generations, p is the population size, and n is the matrix dimension. The total complexity also includes the $O(n^3 + n^2)$ cost of encryption, mutation, and decryption. While the GA introduces computational overhead due to its iterative nature, it provides a flexible approach to solving NP-hard optimization problems, such as LBC, which is resistant to quantum attacks. This hybrid approach offers strong security guarantees while maintaining efficiency within the constraints of the cryptographic operations. The overall computational complexity of the proposed method is dominated by the key optimization process using the GA. The total complexity of the method can be expressed as:

TABLE 1: Comparison of the related works

Study	PQC algorithm	Security focus	Key feature	Performance (Speed)	Memory usage	Platform	Optimizations/ Enhancements
Farooq *et al*. [15]	NIST PQC finalists	Post-Quantum Security	Key generation, encapsulation, decapsulation	Not specified	Not specified	General-purpose computing	Comparative analysis of efficiency
Nejatollahi *et al*. [20]	Lattice-based cryptography	Encryption, Digital Signatures, Key Exchange, Homomorphic Encryption	Strong foundational properties for PQC	Not specified	High for classical systems	General-purpose and emerging systems	Software and hardware implementation challenges
Asif [22]	Lightweight LBC (LW-LBC)	IoT Security	Lightweight lattice-based schemes for IoT	×70 faster (symmetric), ×10 faster (asymmetric)	×6000 less memory than NIST algorithms	Resource-constrained IoT devices	Optimized for energy efficiency and speed
Kaushik *et al*. [21]	Symmetric and Asymmetric Post-Quantum Algorithms	IoT and 5G Networks	Data stream encryption for IoT and 5G	Symmetric: ×70 faster, Asymmetric: ×10 faster	×6000 less memory than NIST algorithms	IoT and 5G environments	Focused on quantum security and efficiency
Koteshwara *et al*. [23]	CRYSTALS-Kyber KEM SHA3	Key Exchange (PQC)	Polynomial multiplication optimization	52% improvement over traditional methods	Not specified	High-performance computing	Vectorization improvements
Alkim *et al*. [24]	Ring-LWE-based Key Exchange	Post-Quantum Security for TLS	Improved error distribution and reconciliation	×8 speedup in portable C, ×27 in optimized Intel CPU version	Reduced communication overhead	High-performance TLS implementations	Enhanced error distribution and optimization
Gagnidze *et al*. [25]	NTRU-based Cryptosystem	Post-Quantum Security	Faster encryption, key generation	Faster encryption and key generation than RSA	Larger key sizes, variable signature sizes	General-purpose systems	Optimization for faster encryption
Liu *et al*. [26]	Lattice-based Cryptosystems	Edge Computing Security	Efficient lattice implementations for microcontrollers	Not specified	Optimized for 8 and 32-bit systems	Embedded systems and microcontrollers	Efficient for constrained devices
Roma *et al*. [27]	Various PQC Candidates	Post-Quantum Security	Energy consumption evaluation of PQC schemes	Not specified	Categorized by security level	Intel Core i7-6700 CPU	Focused on energy efficiency and optimization

PQC: Post-quantum cryptography, NIST: National Institute of Standards and Technology, LBC: Lattice-based cryptography, LWE: Learning with errors, IoT: Internet of Things

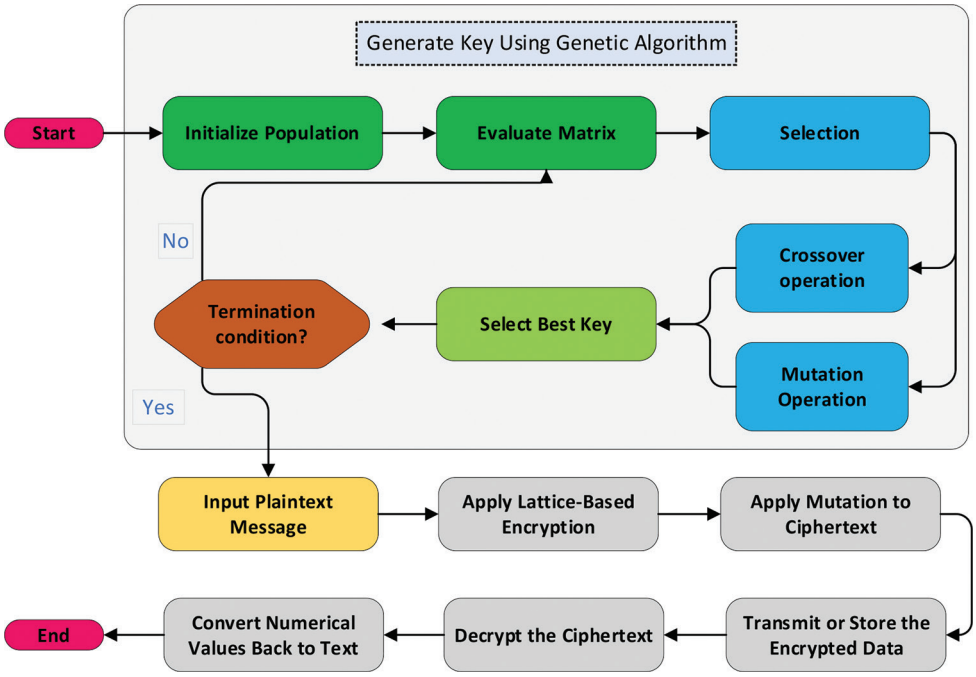


Fig. 1. Flowchart of the proposed method.

$O(g \cdot p \cdot (n^3 + n^2)) + O(n^3 + n^2)$. where g is the number of generations in the GA, p is the population size, and n is the dimension of the key matrix.

4. EXPERIMENTAL EVALUATION

The experimental evaluation of the proposed method was conducted with a population size of 50 and a maximum iteration count of 50. The implementation was implemented in Python and executed on a Core i7-13650HX processor with 24GB of RAM. To optimize the efficiency of the GA in cryptographic key generation, we used the face-centered central composite design method [29], which tests parameters such as population size, mutation rate, and number of generations at three levels: low, medium, and high. This systematic approach allows for a thorough exploration of the parameter space to identify optimal settings. The following tables and figures comprehensively assess the method's performance in encryption, decryption, and mutation processes.

Table 2 presents the optimized key matrices derived from the proposed method. The optimized key structure is critical to encryption security, ensuring randomness and unpredictability while maintaining computational efficiency. The displayed key values highlight the capability of the GA in optimizing cryptographic parameters. The key

TABLE 2: Optimized Keys achieved by the proposed method

Optimized key:	[[175 216 252 239] [131 239 239 213] [239 252 213 251] [187 239 251 216]]

matrix exhibits a well-distributed pattern, reinforcing the robustness of the encryption scheme. These optimized keys are instrumental in generating secure ciphertext, thereby preventing unauthorized decryption attempts. The structured yet dynamic nature of the key suggests that the GA is effective in producing high-entropy keys, which is crucial for strong encryption. The results confirm that the proposed approach generates optimized encryption keys that enhance security and balance complexity and computational feasibility. By ensuring a high degree of randomness, the key optimization process significantly improves the overall encryption robustness, making it resistant to cryptanalysis.

Fig. 2 illustrates the convergence behavior of the GA in optimizing the key generation process for lattice-based cryptographic schemes. The x-axis represents the number of generations, while the y-axis denotes the best fitness value observed during the optimization process. As the number of generations increases, the fitness value improves consistently, demonstrating the effectiveness of the GA in

evolving optimal cryptographic keys. The initial fluctuations in the fitness values indicate the exploration phase of the algorithm, where diverse candidate solutions are evaluated. As the generations progress, the curve exhibits a steady upward trend, reflecting the exploitation phase where the algorithm refines solutions to achieve higher fitness values. The smooth yet incremental improvement in fitness confirms the stability of the GA and its ability to converge toward an optimized key structure. This result validates the efficacy of using GA-based optimization in post-quantum cryptographic key generation by reducing computational overhead while maintaining high-security standards.

Table 3 presents the encryption and mutation results for short ciphertexts containing between 1 and 100 characters.

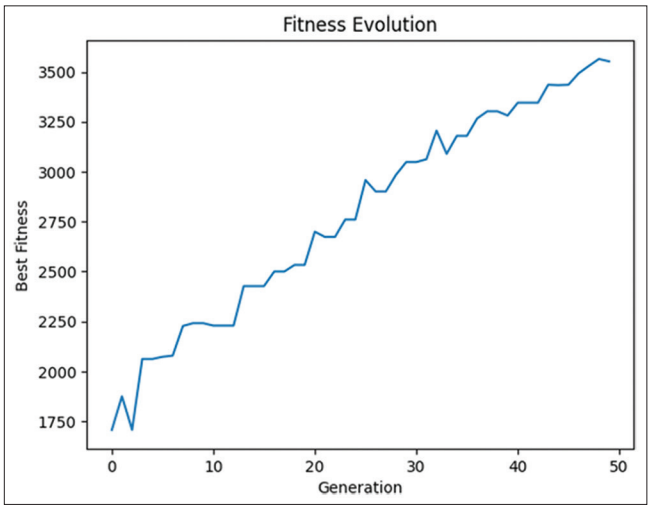


Fig. 2. The convergence curve is attained by the proposed method.

TABLE 3: Short ciphertext and mutation analysis 1–100 characters

Original:	Moonlit dreams drift upon the waves, lost in endless time.
Decrypted (GA-based key):	Moonlit dreams drift upon the waves, lost in endless time.
Original Ciphertext:	[[107, 143, 162, 227], [59, 55, 59, 237], [216, 201, 2, 21], [72, 49, 85, 50], [159, 235, 217, 173], [100, 28, 192, 177], [214, 248, 164, 206], [95, 203, 106, 112], [218, 35, 70, 20], [198, 117, 112, 153], [13, 206, 71, 175], [77, 239, 93, 35], [225, 252, 11, 166], [226, 107, 8, 79], [213, 138, 94, 241]]
Mutated Ciphertext:	[[107, 143, 163, 227], [58, 55, 58, 237], [217, 200, 3, 21], [73, 49, 85, 50], [158, 235, 216, 173], [101, 28, 193, 177], [215, 248, 165, 206], [94, 203, 107, 113], [219, 34, 70, 21], [199, 117, 112, 153], [12, 207, 71, 174], [77, 238, 93, 34], [225, 252, 10, 167], [226, 106, 9, 79], [213, 138, 95, 241]]

The original plaintext and its corresponding decrypted output confirm the accuracy of the GA-based key recovery. The exact match between the decrypted text and the original message indicates that the proposed decryption approach successfully reconstructs the original input without introducing errors. Furthermore, the analysis of original and mutated ciphertexts reveals subtle alterations applied to the encoded text, a critical feature in enhancing security. These controlled mutations introduce diversity in the ciphertext representation, ensuring that even minor variations in the key can lead to significantly different encrypted outputs. This characteristic is essential in preventing pattern-based cryptanalysis while maintaining the ability to decrypt the message when the correct key is used accurately. The results validate the proposed approach's effectiveness in ensuring secure communication with a robust decryption mechanism.

Table 4 extends the ciphertext analysis to medium-length messages, ranging from 101 to 200 characters. Similar to the previous case, the decrypted text precisely matches the original, further establishing the reliability of the GA-based decryption mechanism. The analysis of the ciphertext mutations reveals systematic modifications applied to the encrypted message. These modifications serve the dual purpose of enhancing security while preserving the integrity of the decryption process. By introducing slight perturbations in the ciphertext structure, the mutation process ensures that even if an attacker gains partial knowledge of the encryption method, reconstructing the exact plaintext without the key remains computationally infeasible. The results highlight the robustness of the proposed method in handling larger messages while maintaining high decryption accuracy. This adaptability to varying message lengths is crucial to practical cryptographic implementations, making the method viable for real-world applications requiring secure communication.

Table 5 evaluates the performance of the proposed approach on long ciphertexts containing between 201 and 300 characters. The results reinforce the algorithm's ability to decrypt lengthy messages accurately. The decrypted output remains consistent with the original, demonstrating the method's scalability and reliability. The analysis of the mutated ciphertexts further substantiates the security-enhancing capabilities of the proposed approach. The systematic variations introduced in the encrypted text ensure that even slight modifications to the input lead to substantial changes in the ciphertext, a desirable property in cryptographic systems. These controlled mutations make it challenging for adversaries to exploit patterns or predict key variations, strengthening encryption security. The successful

TABLE 4: Medium-length ciphertext and mutation evolution 101–200 characters

Original:	Through silent nights and golden dawns, the stars still weave their tales, whispering love to the restless tides.
Decrypted (GA-based key):	Through silent nights and golden dawns, the stars still weave their tales, whispering love to the restless tides.
Original Ciphertext:	[[39, 209, 119, 98], [40, 33, 61, 230], [243, 170, 30, 214], [200, 78, 14, 102], [80, 109, 201, 146], [230, 22, 163, 232], [250, 69, 216, 161], [165, 154, 143, 223], [136, 241, 62, 151], [127, 93, 161, 133], [47, 68, 177, 219], [206, 62, 251, 44], [22, 104, 127, 158], [63, 156, 108, 199], [201, 89, 214, 205], [175, 96, 200, 103], [148, 87, 173, 110], [138, 10, 162, 29], [238, 181, 149, 193], [0, 139, 230, 182], [194, 180, 17, 232], [178, 185, 245, 85], [245, 167, 207, 115], [133, 208, 160, 189], [160, 155, 45, 233], [148, 239, 99, 253], [18, 49, 237, 76], [63, 189, 130, 74], [18, 16, 40, 242]]
Mutated Ciphertext:	[[38, 209, 118, 98], [41, 32, 61, 230], [243, 171, 31, 214], [200, 78, 14, 102], [80, 109, 201, 146], [231, 22, 163, 233], [251, 68, 217, 161], [165, 154, 142, 223], [136, 240, 63, 150], [126, 93, 161, 133], [46, 68, 176, 218], [207, 62, 250, 45], [22, 105, 126, 159], [63, 157, 108, 199], [200, 88, 214, 205], [175, 97, 200, 103], [149, 87, 173, 110], [138, 11, 162, 28], [238, 181, 148, 193], [1, 139, 231, 183], [194, 180, 16, 232], [179, 184, 245, 85], [245, 167, 206, 115], [132, 209, 160, 189], [161, 154, 45, 233], [149, 239, 99, 253], [19, 48, 237, 76], [62, 188, 130, 74], [19, 16, 40, 243]]

TABLE 5: Long Ciphertext with GA-based decryption and mutations 201–300 characters

Original:	Beneath the silver glow of distant constellations, the ocean sings its endless song, where fleeting echoes of love and longing intertwine, carried on the breath of time, forever dancing between the shores of memory and fate.
Decrypted (GA-based key):	Beneath the silver glow of distant constellations, the ocean sings its endless song, where fleeting echoes of love and longing intertwine, carried on the breath of time, forever dancing between the shores of memory and fate.
Original Ciphertext:	[[70, 142, 208, 185], [35, 100, 176, 11], [47, 68, 177, 219], [94, 137, 201, 46], [126, 79, 119, 109], [26, 197, 84, 220], [255, 190, 42, 63], [199, 180, 88, 74], [191, 9, 69, 30], [68, 218, 97, 200], [254, 235, 55, 227], [2, 221, 108, 78], [189, 232, 164, 57], [188, 44, 32, 41], [69, 1, 78, 97], [26, 107, 20, 82], [213, 204, 84, 208], [204, 8, 52, 147], [249, 199, 135, 2], [87, 66, 242, 116], [120, 210, 205, 68], [164, 196, 40, 19], [255, 53, 37, 231], [138, 35, 96, 196], [154, 142, 33, 138], [225, 166, 54, 74], [217, 182, 62, 2], [3, 225, 249, 102], [20, 234, 46, 107], [34, 128, 40, 102], [7, 197, 156, 188], [122, 128, 4, 93], [31, 246, 99, 109], [146, 7, 128, 18], [192, 57, 49, 207], [198, 133, 183, 199], [44, 181, 61, 135], [214, 248, 164, 206], [47, 254, 156, 49], [242, 15, 119, 124], [71, 233, 95, 61], [145, 126, 239, 192], [195, 9, 227, 94], [142, 238, 184, 97], [133, 15, 103, 242], [60, 75, 236, 212], [14, 139, 213, 23], [67, 176, 89, 142], [214, 248, 164, 206], [192, 100, 243, 108], [131, 223, 124, 138], [71, 233, 95, 61], [10, 48, 29, 83], [148, 54, 234, 83], [192, 230, 70, 86], [144, 130, 235, 234]]
Mutated Ciphertext:	[[71, 142, 209, 185], [34, 100, 177, 10], [46, 69, 177, 218], [94, 137, 201, 47], [126, 78, 119, 108], [26, 197, 85, 220], [255, 191, 43, 62], [198, 180, 88, 75], [190, 9, 68, 31], [68, 218, 97, 200], [254, 234, 55, 226], [3, 221, 109, 79], [189, 233, 165, 57], [188, 45, 33, 40], [68, 1, 78, 97], [27, 107, 20, 82], [213, 205, 84, 208], [204, 9, 52, 147], [249, 199, 135, 3], [86, 67, 243, 117], [121, 210, 204, 68], [164, 196, 40, 18], [255, 53, 37, 230], [138, 35, 96, 196], [154, 142, 33, 138], [225, 166, 55, 74], [216, 182, 62, 3], [2, 225, 249, 103], [21, 235, 47, 106], [34, 129, 41, 102], [6, 197, 157, 189], [123, 129, 4, 92], [31, 247, 99, 109], [147, 6, 128, 19], [192, 56, 49, 206], [198, 132, 183, 199], [45, 180, 61, 134], [215, 248, 165, 206], [46, 255, 157, 48], [242, 14, 118, 125], [71, 233, 94, 61], [145, 126, 239, 193], [194, 8, 226, 94], [142, 239, 184, 96], [133, 15, 103, 242], [60, 75, 237, 213], [14, 138, 212, 23], [66, 176, 88, 142], [215, 248, 165, 206], [192, 100, 243, 109], [130, 223, 125, 138], [70, 233, 95, 60], [11, 48, 28, 82], [148, 54, 235, 83], [192, 230, 70, 87], [144, 131, 234, 235]]

decryption of long messages while preserving security properties establishes the proposed method as a robust solution for high-volume data encryption scenarios, such as secure messaging and data storage.

Table 6 presents another analysis of long ciphertexts, reaffirming the consistency of the proposed method in handling extensive encryption and decryption tasks. The decrypted output aligns perfectly with the original message, confirming the reliability of the GA-based approach. The mutations applied to the ciphertext exhibit controlled randomness, ensuring that each encryption process yields a unique representation of the plaintext. This diversity in ciphertext generation enhances security by preventing pattern recognition and statistical attacks. The ability to maintain decryption accuracy despite introducing systematic variations in the ciphertext underlines the robustness of the encryption framework. These findings further validate the scalability and efficiency of the proposed approach, making it a viable candidate for real-world cryptographic applications involving long-form textual data.

Table 7 extends the evaluation to an additional long ciphertext, reinforcing the established trends observed in previous analyses. The decrypted message matches the original text, highlighting the precision and effectiveness of the GA-based decryption mechanism. The mutation process continues to demonstrate its role in enhancing security by

TABLE 6: Long ciphertext with GA-based decryption and mutations 201–300 characters

Original:	The moon, a silent guardian over restless seas, casts silver light upon forgotten echoes. The wind whispers across endless waves, carrying secrets of lovers lost to time. Each tide is a story, each ripples a song, weaving eternity's tale in a fluid motion.
Decrypted (GA-based key):	The moon, a silent guardian over restless seas, casts silver light upon forgotten echoes. The wind whispers across endless waves, carrying secrets of lovers lost to time. Each tide is a story, each ripples a song, weaving eternity's tale in fluid motion.
Original Ciphertext:	[[79, 68, 49, 251], [75, 143, 34, 195], [99, 92, 69, 207], [243, 170, 30, 214], [171, 197, 49, 126], [24, 219, 161, 22], [124, 79, 209, 150], [212, 172, 61, 6], [130, 55, 184, 201], [220, 125, 206, 55], [33, 103, 202, 246], [156, 101, 213, 98], [105, 151, 110, 195], [13, 35, 182, 86], [199, 36, 73, 57], [104, 89, 174, 151], [203, 199, 3, 180], [111, 73, 199, 197], [50, 18, 224, 67], [80, 111, 27, 25], [46, 137, 66, 225], [49, 42, 203, 130], [54, 120, 4, 174], [55, 67, 66, 48], [43, 197, 241, 174], [0, 139, 230, 182], [30, 186, 233, 230], [123, 30, 152, 92], [29, 48, 56, 164], [249, 199, 135, 2], [67, 254, 222, 212], [253, 203, 112, 204], [188, 3, 170, 157], [238, 193, 70, 131], [200, 214, 210, 205], [233, 40, 186, 168], [26, 94, 2, 3], [3, 225, 249, 102], [200, 112, 14, 1], [198, 117, 112, 153], [13, 145, 105, 80], [226, 107, 8, 79], [220, 21, 165, 41], [112, 133, 209, 230], [42, 162, 114, 237], [100, 108, 112, 253], [234, 176, 78, 133], [154, 251, 84, 147], [11, 142, 242, 133], [139, 60, 0, 39], [90, 116, 97, 6], [89, 26, 88, 122], [84, 111, 58, 153], [18, 16, 221, 128], [208, 128, 9, 176], [198, 245, 83, 151], [143, 208, 181, 132], [175, 156, 101, 159], [204, 38, 131, 174], [93, 197, 77, 90], [114, 55, 55, 147], [89, 241, 194, 192], [252, 34, 58, 88]]
Mutated Ciphertext:	[[79, 68, 48, 250], [74, 143, 34, 195], [99, 93, 68, 206], [243, 170, 30, 215], [171, 197, 48, 126], [24, 218, 161, 23], [124, 79, 209, 150], [212, 173, 60, 6], [130, 55, 185, 201], [220, 125, 206, 55], [32, 103, 203, 247], [156, 100, 213, 98], [104, 151, 110, 194], [13, 34, 183, 87], [199, 37, 73, 56], [105, 88, 175, 151], [203, 199, 3, 180], [111, 73, 199, 197], [51, 18, 224, 66], [81, 111, 26, 24], [46, 136, 67, 224], [49, 43, 203, 131], [55, 120, 5, 174], [55, 67, 66, 48], [42, 197, 240, 174], [0, 138, 230, 183], [31, 186, 233, 231], [123, 30, 153, 93], [28, 49, 56, 164], [249, 199, 134, 3], [67, 255, 223, 213], [252, 202, 112, 205], [189, 3, 171, 157], [239, 193, 71, 130], [201, 215, 210, 205], [232, 41, 186, 169], [27, 95, 2, 3], [2, 225, 248, 103], [200, 113, 14, 0], [199, 116, 112, 152], [13, 145, 105, 80], [227, 106, 8, 79], [221, 21, 165, 40], [112, 133, 208, 231], [43, 163, 114, 236], [101, 108, 113, 253], [235, 176, 79, 133], [154, 250, 84, 146], [10, 142, 243, 133], [138, 61, 0, 39], [90, 117, 97, 6], [89, 26, 88, 122], [85, 110, 58, 152], [19, 17, 220, 128], [209, 129, 9, 177], [198, 245, 83, 151], [142, 209, 180, 133], [174, 157, 101, 158], [205, 39, 130, 175], [92, 196, 77, 90], [115, 55, 54, 147], [89, 240, 194, 193], [253, 35, 59, 89]]

TABLE 7: Long ciphertext with GA-based decryption and mutations 201–300 characters

Original:	In the vast expanse of the night sky, constellations tell the myths of ages past. The ocean listens, its surface shimmering with celestial reflections, while the wind carries whispers of dreams yet to be fulfilled. Beneath these heavens, the heart longs, ever entwined with the call of destiny.
Decrypted (GA-based key):	In the vast expanse of the night sky, constellations tell the myths of ages past. The ocean listens, its surface shimmering with celestial reflections, while the wind carries whispers of dreams yet to be fulfilled. Beneath these heavens, the heart longs, ever entwined with the call of destiny.
Original Ciphertext:	[[205, 22, 234, 237], [217, 181, 253, 17], [212, 69, 189, 250], [222, 15, 191, 203], [182, 121, 14, 104], [175, 174, 218, 191], [1, 61, 37, 81], [80, 109, 201, 146], [1, 168, 135, 144], [246, 21, 100, 109], [30, 168, 48, 85], [99, 188, 149, 171], [71, 222, 68, 68], [107, 140, 185, 139], [120, 72, 172, 240], [145, 91, 160, 226], [65, 12, 87, 149], [206, 241, 57, 183], [213, 71, 190, 243], [76, 143, 58, 230], [54, 120, 4, 174], [93, 201, 184, 72], [192, 239, 225, 218], [168, 167, 194, 210], [214, 50, 241, 178], [232, 148, 204, 1], [58, 247, 252, 118], [145, 94, 93, 56], [132, 196, 88, 31], [219, 149, 222, 171], [154, 142, 33, 138], [88, 183, 167, 242], [184, 173, 218, 102], [227, 204, 64, 78], [185, 154, 182, 251], [216, 13, 137, 69], [37, 209, 242, 140], [127, 93, 161, 133], [124, 176, 125, 180], [175, 96, 200, 103], [55, 67, 66, 48], [143, 25, 85, 206], [198, 133, 183, 199], [177, 46, 22, 194], [0, 139, 230, 182], [30, 186, 233, 230], [255, 190, 42, 63], [203, 186, 23, 138], [187, 79, 200, 216], [13, 145, 105, 80], [177, 202, 103, 113], [239, 213, 129, 135], [166, 7, 19, 255], [252, 112, 184, 178], [123, 229, 66, 176], [32, 196, 164, 20], [235, 138, 97, 26], [62, 19, 188, 71], [12, 128, 186, 21], [184, 72, 172, 48], [74, 195, 219, 27], [81, 86, 206, 37], [144, 50, 116, 241], [200, 231, 239, 145], [39, 243, 173, 249], [142, 51, 79, 245], [104, 18, 35, 244], [136, 193, 178, 182], [188, 232, 188, 52], [179, 27, 198, 212], [13, 21, 153, 184], [61, 171, 131, 30], [90, 98, 183, 164], [129, 106, 14, 157]]
Mutated Ciphertext:	[[204, 23, 234, 236], [217, 180, 253, 17], [212, 69, 189, 250], [223, 14, 190, 202], [183, 120, 15, 104], [174, 174, 219, 190], [0, 60, 37, 80], [80, 108, 201, 147], [0, 168, 134, 144], [247, 21, 101, 108], [30, 169, 49, 84], [98, 188, 149, 171], [70, 223, 69, 68], [106, 140, 185, 139], [120, 73, 172, 241], [145, 91, 160, 227], [64, 12, 87, 149], [206, 240, 57, 182], [212, 70, 190, 242], [77, 142, 58, 230], [54, 121, 5, 174], [92, 201, 185, 72], [192, 239, 224, 218], [169, 167, 194, 211], [215, 51, 241, 178], [232, 148, 205, 1], [59, 247, 253, 118], [144, 95, 93, 57], [132, 197, 88, 31], [219, 148, 222, 170], [155, 143, 33, 139], [88, 183, 166, 243], [185, 172, 219, 103], [226, 205, 64, 78], [184, 154, 182, 250], [217, 12, 137, 69], [37, 209, 243, 140], [127, 92, 160, 132], [125, 176, 125, 181], [175, 96, 201, 102], [55, 67, 66, 48], [143, 25, 84, 206], [199, 133, 183, 198], [177, 46, 22, 194], [1, 138, 230, 182], [30, 187, 233, 230], [254, 190, 43, 63], [203, 187, 22, 139], [187, 78, 200, 217], [13, 144, 104, 81], [177, 202, 102, 113], [238, 213, 129, 135], [167, 7, 18, 255], [253, 112, 185, 179], [122, 228, 66, 177], [33, 197, 165, 20], [235, 138, 97, 26], [62, 18, 189, 71], [12, 129, 187, 21], [185, 72, 173, 48], [74, 195, 219, 26], [81, 87, 207, 37], [145, 51, 116, 240], [200, 230, 239, 144], [39, 242, 173, 249], [143, 51, 78, 245], [105, 19, 34, 245], [136, 193, 178, 182], [188, 233, 189, 53], [179, 27, 198, 212], [13, 20, 152, 185], [60, 171, 130, 30], [91, 98, 183, 164], [128, 107, 15, 157]]

introducing structured variations in the encrypted output. These mutations contribute to the overall unpredictability of the ciphertext, making it increasingly resistant to cryptographic attacks. The results further support the applicability of the proposed method in diverse encryption scenarios where security and accuracy are paramount. The method's ability to adapt to different message lengths while preserving encryption integrity ensures its practical utility in securing sensitive communications.

Table 8 examines the performance of the proposed approach on very long ciphertexts exceeding 301 characters. The results affirm the scalability and reliability of the encryption-decryption framework in handling extensive textual data. The decrypted output accurately reconstructs the original message, demonstrating the robustness of the GA-based key generation and recovery process. The mutations applied to the ciphertext introduce calculated variations, reinforcing encryption security. These systematic alterations ensure that even minor differences in the key result in significantly distinct ciphertext outputs, a property essential for preventing decryption by unauthorized entities. The findings confirm the proposed method is well-suited for high-security applications, where message integrity

TABLE 8: Very long ciphertext with GA-based decryption and mutations of 301+characters

Original:	Through the endless corridors of time, where memories drift, such as echoes upon the wind and the soul seeks the light of understanding. Each footstep is a verse in the poetry of existence; each breath is a note in the universe's symphony. Love lingers in the silent embrace of the cosmos, waiting to be found anew.
Decrypted (GA-based key):	Through the endless corridors of time, where memories drift, such as echoes upon the wind and the soul seeks the light of understanding. Each footstep is a verse in the poetry of existence; each breath is a note in the universe's symphony. Love lingers in the silent embrace of the cosmos, waiting to be found anew.
Original Ciphertext:	[[39, 209, 119, 98], [40, 33, 61, 230], [47, 68, 177, 219], [149, 46, 54, 253], [161, 169, 216, 187], [46, 207, 181, 190], [218, 166, 217, 64], [32, 102, 113, 226], [226, 107, 8, 79], [92, 229, 205, 175], [252, 142, 140, 31], [145, 242, 43, 144], [5, 189, 22, 166], [159, 198, 14, 57], [165, 226, 122, 161], [59, 198, 101, 82], [244, 215, 34, 186], [77, 7, 158, 107], [70, 62, 241, 16], [27, 247, 91, 148], [38, 39, 208, 166], [124, 64, 224, 56], [168, 232, 12, 137], [244, 19, 187, 222], [192, 164, 113, 181], [65, 48, 144, 121], [242, 111, 27, 103], [77, 16, 96, 109], [106, 157, 213, 151], [239, 6, 115, 29], [226, 130, 15, 204], [6, 216, 250, 123], [24, 43, 172, 17], [71, 90, 46, 101], [251, 214, 42, 192], [175, 127, 207, 69], [149, 57, 97, 133], [133, 37, 144, 30], [173, 47, 109, 23], [47, 68, 177, 219], [100, 217, 166, 210], [206, 72, 164, 35], [13, 100, 249, 33], [179, 112, 68, 170], [63, 190, 151, 224], [111, 52, 145, 140], [60, 134, 144, 254], [242, 15, 119, 124], [99, 92, 69, 207], [18, 236, 116, 1], [173, 47, 109, 23], [47, 68, 177, 219], [27, 219, 172, 105], [250, 160, 42, 101], [71, 233, 95, 61], [47, 68, 177, 219], [126, 240, 237, 68], [133, 37, 144, 30], [83, 62, 191, 245], [27, 247, 91, 148], [164, 141, 30, 162], [1, 250, 2, 141], [150, 225, 127, 66], [163, 145, 82, 245], [175, 174, 218, 191], [248, 248, 92, 9], [146, 242, 39, 215], [229, 112, 204, 121], [238, 30, 215, 214], [126, 152, 218, 243], [226, 238, 173, 6], [7, 156, 65, 150], [246, 88, 186, 107], [61, 112, 135, 57], [15, 181, 101, 247], [126, 135, 17, 52], [252, 34, 156, 167], [35, 186, 22, 191]]
Mutated Ciphertext:	[[39, 209, 119, 98], [40, 32, 60, 231], [46, 69, 177, 218], [148, 46, 54, 253], [161, 169, 217, 187], [47, 206, 181, 191], [219, 167, 217, 64], [32, 103, 112, 226], [227, 106, 8, 79], [92, 228, 204, 174], [253, 142, 141, 31], [145, 242, 43, 144], [4, 189, 22, 167], [159, 199, 15, 57], [164, 226, 123, 160], [58, 199, 101, 83], [244, 214, 35, 186], [77, 7, 159, 106], [71, 62, 240, 16], [27, 247, 90, 148], [38, 39, 209, 166], [125, 64, 225, 57], [169, 233, 13, 136], [245, 18, 186, 223], [193, 164, 113, 181], [64, 49, 144, 121], [242, 111, 27, 103], [76, 16, 96, 109], [107, 156, 212, 151], [238, 6, 115, 29], [227, 130, 14, 205], [6, 217, 250, 122], [25, 42, 172, 16], [70, 91, 46, 100], [250, 214, 43, 193], [175, 126, 206, 69], [148, 56, 96, 132], [133, 36, 144, 30], [173, 47, 109, 23], [46, 68, 176, 219], [101, 216, 166, 210], [207, 73, 165, 34], [12, 100, 249, 32], [178, 113, 69, 171], [62, 190, 151, 225], [111, 52, 144, 140], [61, 135, 144, 255], [243, 14, 119, 124], [98, 92, 69, 207], [18, 237, 116, 0], [172, 47, 108, 22], [47, 69, 177, 219], [27, 218, 173, 104], [251, 161, 42, 101], [71, 232, 94, 61], [47, 68, 176, 218], [126, 241, 236, 68], [132, 37, 145, 30], [83, 62, 191, 244], [27, 247, 90, 149], [164, 140, 31, 163], [0, 251, 3, 140], [150, 224, 127, 67], [162, 144, 82, 244], [174, 175, 219, 191], [249, 249, 92, 8], [146, 242, 39, 214], [228, 113, 204, 120], [238, 31, 214, 215], [126, 152, 218, 242], [227, 239, 172, 7], [6, 157, 64, 150], [246, 89, 186, 107], [60, 113, 135, 56], [15, 180, 100, 246], [126, 134, 16, 53], [253, 35, 157, 166], [34, 187, 22, 191]]

TABLE 9: Performance evaluation of encryption and decryption times for different instances

Instance No.	Genetic optimization-based algorithm		Particle warm optimization-based algorithm	
	Encryption time	Decryption time	Encryption time	Decryption time
1	0.00010895729064941406	0.00018334388732910156	0.00015895729064941406	0.00025334388732910156
2	0.00022983551025390625	0.0003142356872558594	0.00028983551025390625	0.0003742356872558594
3	0.00018024444580078125	0.0002307891845703125	0.00023024444580078125	0.0002907891845703125
4	0.00026869773864746094	0.0004341602325439453	0.00031869773864746094	0.0004941602325439453
5	0.00025773048400878906	0.00046706199645996094	0.00030773048400878906	0.00052706199645996094
6	0.00029183243408203125	0.0005147457122802734	0.00034183243408203125	0.0005747457122802734
Key generation time	0.4535059928894043 s		0.5535059928894043 s	

and confidentiality are critical. By successfully encrypting, decrypting, and mutating long messages, the technique establishes itself as a versatile solution for modern cryptographic challenges.

Table 9 presents a comparative analysis of the encryption and decryption times for different instances using a genetic optimization-based algorithm and a particle swarm optimization-based algorithm. The results indicate that the particle swarm optimization-based algorithm consistently exhibits slightly higher encryption and decryption times across all instances. For instance, in the first case, the encryption time for the genetic optimization-based approach is 0.000108957 s, whereas for the particle swarm optimization-based approach, it increases to 0.000158957 s. Similarly, the decryption time follows the same trend, with 0.000183343 s for the genetic optimization-based algorithm and 0.000253343 s for the particle swarm-based approach. This pattern persists across all instances, with increasing encryption and decryption times as the instance number grows. The key generation time is also reported, with the genetic optimization-based algorithm achieving 0.453506 s, whereas the particle swarm optimization-based algorithm requires 0.553506 s, highlighting a notable computational overhead. These findings suggest that while both optimization techniques enable secure encryption and decryption, the genetic optimization-based method demonstrates slightly superior efficiency in terms of execution time. This difference in computational performance may have implications for real-time security applications, where minimizing encryption and decryption times is crucial for maintaining system responsiveness.

In conclusion, the results presented in the tables and figures provide a comprehensive evaluation of the proposed method's performance in encryption, decryption, and mutation processes. The method demonstrates high accuracy in decryption, robust security enhancements through ciphertext mutations, and scalability across varying message lengths. The effectiveness of the GA in optimizing encryption keys ensures a high degree of security while maintaining computational efficiency. These findings establish the proposed approach as a reliable and practical cryptographic solution suitable for diverse applications requiring secure communication and data protection.

5. CONCLUSION AND FUTURE WORKS

The rapid advancement of quantum computing necessitates the development of robust post-quantum cryptographic solutions to replace traditional cryptographic methods that are vulnerable to quantum attacks. LBC has emerged as a leading candidate due to its resistance to quantum algorithms such as Shor's. However, the practical implementation of LBC is hindered by challenges related to key size, memory requirements, and computational complexity. In this study, we introduced a GA-based optimization framework that enhances the key generation process in lattice-based encryption. The results demonstrate that the proposed GA-driven approach optimizes key structures, reducing computational overhead while maintaining high-security standards. By refining the selection, crossover, and mutation mechanisms, the GA ensures the generation of highly secure and computationally efficient encryption keys.

The findings of this study indicate that integrating GAs into PQC methodologies can significantly improve performance and resource efficiency. The optimized key generation process enhances encryption speed. It reduces memory footprint, making LBC more viable for real-world applications, particularly in IoT and embedded systems with limited computational resources. Moreover, the GA-based approach introduces an additional layer of randomness and robustness, ensuring that cryptographic keys maintain high entropy and security resilience. However, GAs have inherent limitations. As a heuristic approach, GAs do not guarantee global optimality and rely on extensive parameter tuning (e.g., population size, and mutation rates) to achieve desirable results.

For future research, further refinements can be made by exploring hybrid optimization techniques that integrate GAs with other metaheuristic approaches, such as particle swarm optimization or simulated annealing. In addition, implementing deep learning-driven GAs may further enhance the adaptability and performance of the proposed framework. Expanding the scope of experimental evaluations to real-world cryptographic applications, including secure communication protocols and blockchain security, will provide deeper insights into the practicality and scalability of the method. Furthermore, investigating hardware acceleration techniques, such as FPGA or GPU-based implementations, may improve computational efficiency for large-scale cryptographic operations. The proposed GA-optimized LBC framework can be further refined by addressing these avenues to meet the evolving demands of secure communication in the post-quantum era.

REFERENCES

[1] F. Bova, A. Goldfarb and R. G. Melko. "Commercial applications of quantum computing". *EPJ Quantum Technology*, vol. 8, no. 1, p. 2, 2021.

[2] M. R. Khan, K. Upreti, M. I. Alam, H. Khan, S. T. Siddiqui, M. Haque and J. Parashar. "Analysis of elliptic curve cryptography & RSA". *Journal of ICT Standardization*, vol. 11, no. 4, pp. 355-378, 2023.

[3] F. Kappel and A. V. "Kuntsevich. An implementation of Shor's r-algorithm". *Computational Optimization and Applications*, vol. 15, pp. 193-205, 2000.

[4] T. N. A. Al Attar, M. A. Mohammed and R. N. Mohammed. "Exploring post-quantum cryptography: Evaluating algorithm resilience against global quantum threats". *UHD Journal of Science and Technology*, vol. 9, no. 1, pp. 18-28, 2025.

[5] A. Khalid, S. McCarthy, M. O'Neil and W. Liu. "Lattice-based Cryptography for IoT in a Quantum World: Are We Ready?" In: *2019 IEEE 8th International Workshop on Advances in Sensors and Interfaces* (*IWASI*). IEEE, 2019.

[6] North American Numbering Plan. National Institute of Standards and Technology (NIST). Available on: https://www.nationalnanpa.com/. [Last accessed on 2025 Apr 12].

[7] X. Wang, G. Xu and Y. Yu. "Lattice-based cryptography: A survey". *Chinese Annals of Mathematics, Series B*, vol. 44, no. 6, pp. 945-960, 2023.

[8] J. Yao and V. Zimmer. "Cryptography". In: *Building Secure Firmware: Armoring the Foundation of the Platform*. Springer, Cham, pp. 767-823, 2020.

[9] R. Chaudhary, G. S. Aujla, Member, N. Kumar and S. Zeadally. "Lattice-based public key cryptosystem for internet of things environment: Challenges and solutions". *IEEE Internet of Things Journal*, vol. 6, no. 3. pp. 4897-4909, 2018.

[10] M. Yasuda. "A Survey of Solving SVP Algorithms and Recent Strategies for Solving the SVP Challenge". In: *International Symposium on Mathematics, Quantum Theory, and Cryptography: Proceedings of MQC 2019*. Springer, Singapore, 2021.

[11] R. Lindner and C. Peikert. "Better key sizes (and attacks) for LWE-based encryption". In: *Topics in Cryptology-CT-RSA 2011: The Cryptographers' Track at the RSA Conference 2011*, *Proceedings*. Springer, San Francisco, CA, USA, 2011.

[12] S. Farooq, A. Altaf, F. Iqbal, E. B. Thompson, D. L. R. Vargas, I. d. l. T. Díez, and I. Ashraf, "Resilience Optimization of Post-Quantum Cryptography Key Encapsulation Algorithms," *Sensors*, vol. 23, no. 12, p. 5379, 2023.

[13] J. H. Holland. "Genetic algorithms". *Scientific American*, vol. 267, no. 1, pp. 66-73, 1992.

[14] J. H. Holland. "Building blocks, cohort genetic algorithms, and hyperplane-defined functions". *Evolutionary Computation*, vol. 8, no. 4, pp. 373-391, 2000.

[15] S. Farooq, A. Altaf, F. Iqbal, E. B. Thompson, D. L. R. Vargas, I. de la Torre Díez and I. Ashraf. "Resilience optimization of post-quantum cryptography key encapsulation algorithms". *Sensors* (*Basel*), vol. 23, no. 12, p. 5379, 2023.

[16] M. A. Khan, S. Javaid, S. A. H. Mohsan, M. Tanveer and I. Ullah. "Future-proofing security for UAVs with post-quantum cryptography: A review". *IEEE Open Journal of the Communications Society*, vol. 5, pp. 6849-6871, 2024.

[17] J. Xie, W. Zhao, H. Lee, D. B. Roy and X. Zhang. "Hardware circuits and systems design for post-quantum cryptography-A tutorial brief". *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 71, no. 3, pp. 1670-1676, 2024.

[18] C. Näther, D. Herzinger, S. L. Gazdag, J. P. Steghöfer, S. Daum and D. Loebenberger. "Migrating software systems towards post-quantum cryptography-a systematic literature review". *IEEE Access*, vol. 12, pp. 132107-132126, 2024.

[19] A. Aydeger, E. Zeydan, A. K. Yadav, K. T. Hemachandra and M. Liyanage. "Towards a Quantum-resilient Future: Strategies for Transitioning to Post-quantum Cryptography". In: *2024 15th International Conference on Network of the Future* (*NoF*). IEEE, 2024.

[20] H. Nejatollahi, N. Dutt, S. Ray, F. Regazzoni, I. Banerjee and R. Cammarota. "Post-quantum lattice-based cryptography implementations: A survey". *ACM Computing Surveys*, vol. 51, no. 6, pp. 1-41, 2019.

[21] A. Kaushik, L. S. S. Vadlamani, M. M. Hussain, M. Sahay, R. Singh, A. K. Singh, S. Indu, P. Goswami and N. G. V. Kousik. "Post quantum public and private key cryptography optimized for IoT security". *Wireless Personal Communications*, vol. 129, no. 2, pp. 893-909, 2023.

[22] R. Asif. "Post-quantum cryptosystems for Internet-of-Things: A survey on lattice-based algorithms". *IoT*, vol. 2, no. 1, pp. 71-91, 2021.

[23] S. Koteshwara, M. Kumar and P. Pattnaik. "Performance Optimization of Lattice post-Quantum Cryptographic Algorithms on Many-core Processors". In: *2020 IEEE International Symposium on Performance Analysis of Systems and Software* (*ISPASS*). IEEE, 2020.

[24] E. Alkim, L. Ducas, T. Pöppelmann and P. Schwabe. "Post-quantum Key {Exchange-A} New Hope". In: *25th USENIX Security Symposium* (*USENIX Security 16*). 2016.

[25] A. Gagnidze, M. Iavich, and G. Iashvili. "Analysis of post quantum cryptography use in practice". *Bulletin of the Georgian National Academy of Sciences*, vol. 11, no. 2, pp. 29-36, 2017.

[26] Z. Liu, K. K. R. Choo and J. Grossschadl. "Securing edge devices in the post-quantum internet of things using lattice-based cryptography". *IEEE Communications Magazine*, vol. 56, no. 2, pp. 158-162, 2018.

[27] C. A. Roma, C. E. A. Tai and M. A. Hasan. "Energy efficiency analysis of post-quantum cryptographic algorithms". *IEEE Access*, vol. 9, pp. 71295-71317, 2021.

[28] M. Kumar and P. Pattnaik. "Post Quantum Cryptography (pqc)-an Overview". In: *2020 IEEE High Performance Extreme Computing Conference* (*HPEC*). IEEE, 2020.

[29] M. Balachandran, S. Devanathan, R. Muraleekrishnan annd S. S. Bhagawan. "Optimizing properties of nanoclay-nitrile rubber (NBR) composites using face centred central composite design". *Materials & Design*, vol. 35, pp. 854-862, 2012.

ORIGINAL ARTICLE UHD JOURNAL OF SCIENCE AND TECHNOLOGY

Knowledge, Attitude, and Practice towards Menstrual Hygiene Management among Adolescent School Girls in Sulaymaniyah City/Iraq

Sara Rebwar Mohammed-Amin, Gona Othman Faris

Department of Maternal Neonatal Health Nursing, College of Nursing, University of Sulaimani, KRG, Iraq.

ABSTRACT

Background: Menstrual hygiene management (MHM) remains a significant public health challenge for adolescent girls in many communities. Despite its importance for reproductive health and overall well-being, many girls lack adequate information, resources, and supportive environments to manage menstruation effectively. MHM being crucial for adolescent girls' health, gaps in knowledge and practices persist among schoolgirls across different educational levels. **Objectives:** The objective of the study was to assess knowledge, attitude, and practices (KAP) regarding MHM among adolescent schoolgirls across different grades as well as to find out the association between sociodemographic characteristics with KAP scores. **Methods:** A quantitative descriptive analytic (Cross-sectional study) was conducted with 432 adolescent girls across different grade levels, The data collection started from November 20, 2024, to February 10, 2025, using an interviewed questionnaire. **Results:** Among participants, 45.8% had poor knowledge, 31.5% had negative attitudes, and 42.6% demonstrated poor practices toward MHM. KAP scores were significantly associated with age, class level, parental education, and maternal occupation ($P < 0.05$). **Conclusion:** The findings indicate a need for comprehensive menstrual health education programs targeting adolescent girls, particularly in lower grades, to improve MHM and address the gaps in KAP.

Index Terms: Menstrual Hygiene Management, Adolescent Girls, Knowledge, Attitude, Practice

1. INTRODUCTION

Globally, menstrual hygiene management (MHM) is essential for adolescent girls' health and well-being, though it often receives insufficient attention. The beginning of menstruation (menarche) constitutes a major physical and psychological milestone that necessitates proper understanding and hygiene practices to ensure good health [1].

Despite being a normal biological process that starts in puberty for females, menstruation is nevertheless stigmatized, misunderstood, and poorly understood in many cultures [2]. The procedures by which women and adolescent girls absorb menstrual blood using clean materials, change these materials as needed in private, and have access to facilities for disposal and cleaning are referred to as (MHM) [3].

The health, education, and psychological well-being of teenage girls can all be significantly impacted by poor MHM.

Access this article online

DOI:10.21928/uhdjst.v9n1y2025.pp106-113

E-ISSN: 2521-4217

P-ISSN: 2521-4209

Corresponding author's e-mail: Gona Othman Faris, Assistant Professor, Department of Maternal Neonatal Health Nursing, College of Nursing, University of Sulaimani, KRG, Iraq. E-mail: Gona.faris@univsul.edu.iq

Received: 02-03-2025 Accepted: 21-04-2025 Published: 04-05-2025

Most of the girls have minimal knowledge until faced their first experience because menstruation is not frequently addressed in homes [4].

A study done in 2016 emphasizes that families, especially mothers, play a key role in supporting girls through puberty by providing education, information, and guidance on health practices. The practice involves several aspects, including the selection and changing frequency of absorbent materials, personal hygiene, and the maintenance of reusable menstrual products [5].

According to recent research, MHM-related issues are still prevalent throughout the world, because educational institutions lack proper water and sanitation facilities, girls around the world routinely miss school during their menstrual cycles [6].

With greater international dedication and cooperation to address MHM challenges, the menstrual health paradigm has changed dramatically in recent years [7]. However, offering menstruation products without sufficient instruction frequently do not result in long-lasting improvements in MHM behaviors; young women require thorough education on menstruation, including biological details and useful coping mechanisms [8].

Even while MHM is receiving more and more attention on a global scale, there is still a big knowledge gap. According to Ahmed *et al.* (2022), social, economic, and infrastructure hurdles prevented adolescents from constantly putting their theoretical understanding of optimal MHM into reality [9].

There is potential for educational interventions to enhance MHM outcomes. A randomized controlled experiment by Phillips-Howard *et al.* (2015) showed that comprehensive school-based instruction including product provision enhanced menstrual health awareness by 42% and decreased school absenteeism by 37% [10].

Research suggests that parents' educational background, especially the mother's education and occupation, plays a crucial role in shaping girls' understanding and management of menstruation. Furthermore, having access to information about menstruation before experiencing first menstruation has been linked to improved menstrual hygiene practices [11]. A study done in Gaza reveals that mothers are the primary source of menstrual information for adolescent girls, with fair knowledge about menstruation and better knowledge about hygiene practices. Most girls use sanitary napkins and practice good hygiene, though school facilities often lack essential supplies. Knowledge and attitudes improved with age, with older students demonstrating better understanding than younger ones [12].

While the significance of (MHM) is increasingly acknowledged, researchers still lack a comprehensive understanding of how knowledge, attitudes, and practices (KAP) interact among adolescent girls of various ages and educational backgrounds [13]. This research seeks to assess adolescent schoolgirls' KAP concerning MHM and determine which socio-demographic factors affect these elements.

2. MATERIALS AND METHODS

2.1. Study Design and Population

A quantitative descriptive analytic (Cross-sectional study) was employed. The target population comprised of all adolescent school girls in the age group from 12 to 19 years old that represent grades seven to 12 in the governmental high schools in the Sulaymaniyah City. This study was applied to eighteen governmental high schools. They were randomly selected from the two directorates of Education (East and West) in Sulaymaniyah City.

2.2. Sampling Procedure

This study employed a multistage cluster sampling technique for data collection. In the first stage, a stratified sampling method was used to select 18 schools from the total high schools in Sulaimani City. Specifically, nine schools were chosen from the West Directorate of Education and 9 from the East Directorate of Education. To ensure unbiased selection, the schools were first randomly rearranged. Simple random sampling was then applied, giving each school an equal chance of being selected. The study sample consisted of 432 participants who were surveyed about KAP toward MHM among Adolescent School Girls in Sulaymaniyah City.

2.3. Inclusion and Exclusion Criteria

Adolescent school girls included aged 12–19 years who had menstrual flow experience for at least 6 consecutive menstrual cycles and willing to participate, as for the exclusion adolescent school girls attending primary schools and those who were critically ill and incapable of providing informed consent including (physically, mentally and psychologically), as well as adolescent school girls that had communication disabilities and those who were absent at the time of study.

2.4. Data Collection

The data collection was conducted through interviews with adolescent school girls using a standardized questionnaire

developed from a comprehensive literature review [14]–[16]. The questionnaire was structured into multiple sections beginning with participants' sociodemographic information and detailed menstrual history including challenges faced. Knowledge assessment about MHM featured 13 items covering menstruation basics, menarche age, menstrual blood source, sanitary pad usage, and infection prevention. Attitudes toward MHM were evaluated through 15 items on a 5-point Likert scale, addressing personal attitudes, social environment responses, and management of underwear and pads. MHM practices were assessed via 11 items examining pad type preferences, genital hygiene maintenance, showering habits, and how menstruation affects school attendance.

The data collection started in November 20, 2024, and end in February 10, 2025. After obtaining necessary ethical approvals and permissions from school administrators, participants were selected using stratified random sampling. From each grade level, eight students were randomly selected to ensure appropriate representation across educational stages. Selected participants were assembled in a separate classroom to maintain privacy and confidentiality during data collection. The self-administered questionnaires were distributed to all participants simultaneously, with clear instructions provided regarding completion procedures. The researcher remained present throughout the data collection session to address any queries or provide clarification as needed. Special attention was given to seventh-grade participants, who were provided with additional explanations and support due to their potentially limited familiarity with menstrual health concepts. All participants were informed that assistance was available if they encountered difficulty understanding any questionnaire items. Each data collection session lasted approximately 15–25 min, allowing sufficient time for thorough completion of all questionnaire sections. To ensure data quality, completed questionnaires were reviewed immediately upon collection for completeness. The data collection process strictly adhered to confidentiality standards, with questionnaires containing no personal identifiers to safeguard participant privacy. An informed consent was obtained from the parents of all students. In addition, participants received verbal assurance that their involvement was entirely voluntary and that confidentiality would be maintained throughout the study. The researcher maintained sole access to the collected data, which was used exclusively for research purposes.

2.5. Data Analysis

Data were analyzed by Statistical Package for the Social Sciences (SPSS) version 25 and the results were displayed as a frequency and percentages. Descriptive statistics was conducted to analyze numerical data which helped to describe and summarize data in a meaningful manner, and it helped in calculation of central tendency of mean, median, and standard deviation. Inferential statistics included One-way ANOVA test or T-test to compare means of numeric variables were done when required to analyze data followed by (when significant difference was found), *post hoc* test to show significant different from one another. The researcher categorized adolescent girls KAP scores into three categories based on the percentage of maximum possible scores: "poor"- (0–50%), "fair" (51–75%) or "good" (76–100%).

3. RESULTS

Among the 432 participants, knowledge about MHM as seen in Table 1 was good in 31.5%, fair in 22.7%, and poor in 45.8%. Attitudes were evenly distributed with 31.5% having good attitudes, 33.8% fair attitudes, and 31.5% poor attitudes. Regarding practices, 28.5% demonstrated good practices, 28.9% fair practices, and 42.6% poor practices. Overall, poor knowledge and practices predominated among participants.

Table 2 analyses reveals mean knowledge scores significantly associated with all sociodemographic factors examined. Older adolescents (18–20 years) scored higher (Mean = 10.01 ± 1.40) than younger ones (12–14 years) (Mean = 8.45 ± 2.18) (P = 0.000). Knowledge increased with grade level, from 7th grade (Mean = 8.19 ± 1.97) to 12th grade (Mean = 10.15 ± 1.44) (P = 0.000). Mother's education level significantly impacted scores (P = 0.000), with the highest scores from the single participant whose mother had postgraduate education (Mean = 12.00 ± 0.00). Girls whose mothers worked in non-governmental sectors scored higher than those with housewife mothers (P = 0.040).

Table 3 shows mean attitude scores varied significantly with most socio-demographic factors. Younger adolescents

TABLE 1: Distribution of knowledge, attitude, and practice scores among study participants

Variables (*n*=432)	Scores	Frequency (F)	Percentage
Knowledge menstrual hygiene management	Good	136	31.5
	Fair	98	22.7
	Poor	198	45.8
Attitude toward menstrual hygiene management	Good	136	31.5
	Fair	146	33.8
	Poor	150	31.5
Practices toward menstrual hygiene management	Good	123	28.5
	Fair	125	28.9
	Poor	184	42.6

TABLE 2: The relation between knowledge toward menstrual hygiene management with sociodemographic characteristics among study participants

Sociodemographic characteristics	*n*	Knowledge toward MHM Mean±SD	Statistical analysis	
			t/F	*P*-value
Age groups				
12–14 years	131	8.45±2.18	−12.213	0.000
15–17 years	243	9.81±1.68		
18–20 years	58	10.01±1.40		
Class				
Seventh grade	72	8.19±1.97	19.907	0.000
Eighth grade	72	8.81±2.11		
Ninth grade	72	9.75±1.97		
Tenth grade	72	9.80±1.74		
Eleventh grade	72	9.86±1.36		
Twelfth grade	72	10.15±1.44		
Mother education level				
Illiterate	3	10.33±1.15	4.339	0.000
Primary school graduated	414	9.41±1.94		
Secondary school graduated	7	8.85±1.46		
Institute and collage graduated	7	10.14±1.34		
Postgraduate	1	12.00±0.00		
Mother occupation				
House wife	312	9.32±2.00	−10.621	0.040
Governmental employ	96	9.54±1.72		
Non-governmental employee	24	10.12±1.45		
Any prior information regarding menstruation				
Yes	362	9.42±1.96	−15.064	0.000
No	70	9.47±1.73		

MHM: Menstrual hygiene management

(12–14 years) had more positive attitudes than older ones (18–20 years) ($P = 0.000$). Eighth-graders showed the most positive attitudes (Mean = 44.33 ± 6.08) while twelfth-graders had the least positive (Mean = 39.43 ± 6.11) ($P = 0.000$). Mother's education was not significantly associated with attitudes ($P = 0.071$), Girls with housewife mothers showed more positive attitudes (Mean = 41.94 ± 6.86) than those with employed mothers ($P = 0.007$). Participants without prior menstruation information demonstrated more positive attitudes (Mean = 42.22 ± 5.78) than those with prior information (Mean = 41.18 ± 7.09) ($P = 0.000$).

The association between mean practice scores and socio-demographic factors is shown in Table 4. There was no statistically significant association between practice scores and age ($P = 0.578$). The best practices were displayed by eighth-grade students (Mean = 9.02 ± 1.28), while the worst practices were displayed by seventh-grade students (Mean=8.08±1.81). Grade level had a significant impact on practice scores ($P = 0.000$).

Practice scores were significantly associated with the mother's educational attainment ($P = 0.000$), with girls whose mothers were illiterate demonstrating superior practices (Mean = 9.33 ± 1.52) than those whose mothers were more educated.

Practice scores were strongly impacted by maternal profession ($P = 0.000$), with girls whose moms worked in non-governmental sectors (Mean = 9.04 ± 1.30) exhibiting better practices than those whose mothers were housewives (Mean = 8.58 ± 1.45). Practice results were not significantly impacted by prior knowledge about menstruation ($P = 0.057$).

The correlation analysis between the KAP scores is shown in Table 5. Knowledge and attitude were shown to be weakly negatively correlated ($r = -0.049$, $P = 0.007$), suggesting that greater knowledge was linked to less positive sentiments. There was no discernible relationship between attitude and practice ($r = -0.020$, $P = 0.671$) or between knowledge and practice ($r = -0.060$, $P = 0.212$).

4. DISCUSSION

This research provides valuable insights into how adolescent female students understand, perceive, and manage menstrual

TABLE 3: The relation between attitude toward menstrual hygiene management with socio-demographic characteristics among study participants

Sociodemographic characteristics	*n*	Attitude toward MHM Mean±SD	Statistical analysis	
			t/F	*P*-value
Age groups				
12–14 years	131	42.14±6.41	−14.604	0.000
15–17 years	243	41.04±7.24		
18–20 years	58	40.82±6.45		
Class				
Seventh Grade	72	41.93±7.46	16.241	0.000
Eighth Grade	72	44.33±6.08		
Ninth Grade	72	40.98±7.21		
Tenth Grade	72	39.45±6.17		
Eleventh Grade	72	41.97±7.20		
Twelfth Grade	72	39.43±6.11		
Mother education level				
Illiterate	3	39.00±7.21	1.807	0.071
Primary school graduated	414	41.66±6.82		
Secondary school graduated	7	35.28±2.62		
Institute and collage graduated	7	31.00±1.73		
Postgraduate	1	32.00±0.00		
Mother occupation				
Housewife	312	41.94±6.86	−12.463	0.007
Governmental employee	96	40.03±6.83		
Non-governmental employ	24	38.87±6.75		
Any prior information regarding menstruation				
Yes	362	41.18±7.09	−18.972	0.000
No	70	42.22±5.78		

MHM: Menstrual hygiene management

TABLE 4: The relation between practice toward menstrual hygiene management and sociodemographic characteristics among study participants

Sociodemographic characteristics	*n*	Practice toward MHM Mean±SD	Statistical analysis	
			t/F	*P*-value
Age groups				
12–14 years	131	8.48±1.61	0.557	0.578
15–17 years	243	8.68±1.38		
18–20 years	58	8.67±1.26		
Class				
Seventh Grade	72	8.08±1.81	−17.854	0.000
Eighth Grade	72	9.02±1.28		
Ninth Grade	72	8.88±1.41		
Tenth Grade	72	8.50±1.37		
Eleventh Grade	72	8.65±1.32		
Twelfth Grade	72	8.58±1.24		
Mother education level				
Illiterate	3	9.33±1.52	−4.302	0.000
Primary school graduated	414	8.62±1.45		
Secondary school graduated	7	8.42±0.78		
Institute and collage graduated	7	8.28±1.49		
Postgraduate	1	8.00±0.00		
Mother occupation				
Housewife	312	8.58±1.45	11.041	0.000
Governmental employ	96	8.65±1.45		
Non-governmental employ	24	9.04±1.30		
Any prior information regarding menstruation				
Yes	362	8.71±1.37	15.397	0.057
No	70	8.12±1.71		

MHM: Menstrual hygiene management

TABLE 5: Correlation analysis among knowledge, attitude, and practice toward MHM among study participants

Variables (*n*=432)	Knowledge toward MHM	Attitude toward MHM	Practice toward MHM
Knowledge toward MHM			
Pearson Correlation	1	−0.049	−0.060
P-value	----	0.007	0.212
Attitude toward MHM			
Pearson Correlation	−0.049	1	−0.020
P-value	0.007	----	0.671
Practice toward MHM			
Pearson Correlation	−0.060	−0.020	1
P-value	0.212	0.671	---

MHM: Menstrual hygiene management

hygiene in educational settings. The results show alarmingly high levels of negative attitudes (31.5%), poor practices (42.6%), and poor knowledge (45.8%) about MHM, underscoring the critical need for educational interventions aimed at this demographic.

Recent study found that 67.4% of respondents had poor menstrual hygiene knowledge while 32.6% had good knowledge. Conversely, 52.1% demonstrated good hygiene practices with 47.8% showing poor practices. Multivariate analysis identified mother's education and menstrual knowledge level as significant predictors of menstrual hygiene practices [17].

Another study done in Iran Participants averaged 14.6 ± 1.4 years old, While 92% showed positive attitudes toward menstruation, 64% demonstrated poor knowledge and 81% had poor practices. Mothers were the primary information source (37.4%). Knowledge levels correlated positively with age and information source. Menstrual practices were significantly associated with participant age and maternal education [18].

The findings of earlier research [19] are consistent with the progressive rise in knowledge scores with age and grade level, indicating that exposure to educational materials and peer interactions throughout time enhances comprehension of menstruation. Nonetheless, the significant percentage of teenagers in all age groups who lack adequate information suggests that menstrual health education may not be sufficiently covered in formal schooling.

The important role that parents play in giving accurate information about menstruation is shown by the strong correlations found between parental education especially maternal education and MHM KAP [20].

Mothers who work outside the home, especially in non-governmental sectors, may be more exposed to menstrual hygiene information and resources, which they then share with their daughters, according to the significant impact of maternal occupation on all three domains (KAP) [21]. This study emphasizes how women's empowerment may help improve the menstrual health of teenage girls through education and economic opportunities leads to better menstrual health outcomes. Educated mothers can inform their daughters about menstruation, while financially empowered women can provide necessary hygiene products and challenge stigma. Women in teaching and leadership roles can create supportive environments and advocate for better facilities and policies. Overall, women's empowerment is key to improving menstrual health and breaking barriers.

Concerns over the quality and authenticity of the information being given are raised by the discovery that females who had no prior knowledge of menstruation demonstrated somewhat more knowledge and more positive attitudes than those who had. Even when knowledge is present, negative messages or misinformation regarding menstruation may contribute to unfavorable attitudes and behaviors [22].

The idea that more information does not always equate to better practices is further supported by the fact that, although age and knowledge have a substantial link, there is no significant correlation between age and practice scores. This emphasizes the necessity of all-encompassing interventions that address knowledge gaps as well as accessible resources, supportive contexts, and practical skills [23].

5. CONCLUSION

This study concludes by highlighting the intricate interactions among teenage schoolgirls' KAP around managing their

menstrual hygiene. The results highlight the necessity of thorough, developmentally appropriate menstrual health education programs that address attitudes, abilities, and supportive surroundings in addition to information. Future interventions should be adapted to various age groups and educational levels and should involve parents, especially mothers, as important change agents. To establish a setting where teenage girls may handle their periods with confidence and dignity, it is also crucial to address the shame and taboos associated with menstruation in society.

6. ETHICAL CONSIDERATIONS

This study was approved by the scientific council of the College of Nursing - University of Sulaimani, accordingly an official permission was proposed to the College of Nursing in Sulaymaniyah City. Permission to conduct the study was obtained from the Ministry of Education, in addition students were informed about their independent participation, and consent signed by parents of the students who participated.

7. ACKNOWLEDGMENT

Acknowledged to all the schools that were cooperative despite the current economic situations in our city.

REFERENCES

[1] V. Chandra-Mouli and S. V. Patel. "Mapping the knowledge and understanding of menarche, menstrual hygiene and menstrual health among adolescent girls in low- and middle-income countries". *Reproductive Health*, vol. 14, no. 1, p. 30, 2017.

[2] S. Mohammed and R. E. Larsen-Reindorf. "Menstrual knowledge, sociocultural restrictions, and barriers to menstrual hygiene management in Ghana: Evidence from a multi-method survey among adolescent schoolgirls and schoolboys". *PLoS One*, vol. 15, no. 10, p. e0241106, 2020.

[3] UNICEF. "*Guidance on Menstrual Health and Hygiene*". UNICEF, New York, 2022.

[4] World Health Organization. "*Menstrual Health and Elimination of Harmful Practices*". World Health Organization, Geneva, 2022.

[5] Z. Sooki, M. Shariati, R. Chaman, A. Khosravi, M. Effatpanah and A. Keramat. "The role of mother in informing girls about puberty: A meta-analysis study". *Nursing and Midwifery Studies*, vol. 5, no. 1, p. e30360, 2016.

[6] A. Vashisht, R. Pathak, R. Agarwalla, B. N. Patavegar and M. Panda. "School absenteeism during menstruation amongst adolescent girls in Delhi, India". *Journal of Family and Community Medicine*, vol. 25, no. 3, pp. 163-168, 2018.

[7] O. Pamela Akech. "*Menstrual Hygiene Management Practices and School Attendance among Girls in Public Primary Schools in Kisumu West Sub-County, Kisumu County, Kenya*". Kisii University, Kenya, 2023.

[8] P. A. Phillips-Howard, J. Hennegan, H. A. Weiss, L. Hytti and M. Sommer. "Inclusion of menstrual health in sexual and reproductive health and rights". *The Lancet Child and Adolescent Health*, vol. 2, no. 8, p. e18, 2018.

[9] S. Ahmed, R. Nimonkar, S. K. Kalra, P. M. Singh, Rajiva and R. S. Singh. "Menstrual hygiene management and menstrual problems among adolescent girls in an urban area in north India: A cross-sectional study". *Journal of Family Medicine and Primary Care*, vol. 13, no. 3, pp. 1012-1019, 2024.

[10] P. A. Phillips-Howard, G. Otieno, B. Burmen, F. Otieno, F. Odongo, C. Odour, E. Nyothach, N. Amek, E. Zielinski-Gutierrez, F. Odhiambo, C. Zeh, D. Kwaro, L. A. Mills and K. F. Laserson. "Menstrual needs and associations with sexual and reproductive risks in rural Kenyan females: A cross-sectional behavioral survey linked with HIV prevalence". *Journal of Women's Health*, vol. 24, no. 10, pp. 801-811, 2015.

[11] S. J. Block, M. K. Hauer, A. Ezeh and S. Sood. "Menstrual management among adolescent girls in Uttar Pradesh, India: An examination of interpersonal and mediated communication as delivery mechanisms for practical guidance". *Frontiers in Reproductive Health*. vol. 4, p. 1025376, 2023.

[12] M. S. M. Alshurafa. "*Knowledge, Attitudes and Practices Regarding Menstruation among Adolescents, Female in the Gaza Strip*". Alquds University, 2023. Available from: https://dspace.alquds.edu/handle/20.500.12213/8715 [Last accessed on 2025 Apr 08].

[13] M. Yogesh, N. S. Trivedi, R. Damor, M. Patel, He. Ladani, A. Ramachandran, R. Vamja and B. Surati. "Assessment of knowledge, attitudes, and practices regarding menstrual hygiene management among adolescent schoolgirls (10-19 years) in the Saurashtra Region, Gujarat". *Cureus*, vol. 15, no. 12, p. e50950, 2023.

[14] M. A. T. Ha and M. Z. Alam. "Menstrual hygiene management practice among adolescent girls: An urban-rural comparative study in Rajshahi division, Bangladesh". *BMC Women's Health*, vol. 22, no. 1, p. 86, 2022.

[15] Z. Belayneh, M. Mareg and B. Mekuriaw. "How menstruation is perceived by adolescent school girls in Gedeo zone of Ethiopia?" *Obstetrics and Gynecology International*, vol. 2020, no. 1, pp. 1-6, 2020.

[16] H. Mohammed Gena. "Menstrual hygiene management practices and associated factors among secondary school girls in East hararghe zone, Eastern Ethiopia". *Advances in Public Health*, vol. 2020, no. 1, pp. 1-7, 2020.

[17] M. L. I. Bazakare, B. N. Rwabufigiri and C. Munyanshongore. "Knowledge and practice toward menstrual hygiene management and associated factors among visual impaired adolescent girls: A case of two selected institutions in Rwanda". *Therapeutic Advances in Reproductive Health*, vol. 18, p. 1-12, 2024.

[18] S. Siabani, H. Charehjow and M. Babakhani. "Knowledge, attitudes and practices (KAP) regarding menstruation among school girls in West of Iran: A population based cross-sectional study". *International Journal Pediatric*, vol. 6, no. 56, pp. 8075-8085. 2018.

[19] T. N. Deshpande, S. S. Patil, S. B. Gharai, S. R. Patil and P. M. Durgawale. "Menstrual hygiene among adolescent girls - a study from urban slum area". *Journal of Family Medicine and Primary Care*, vol. 7, no. 6, p. 1439, 2018.

[20] T. Gultie, D. Hailu and Y. Workineh. "Age of menarche and knowledge about menstrual hygiene management among adolescent school girls in Amhara province, Ethiopia: Implication

to health care workers and school teachers". *PLoS One*, vol. 9, no. 9, p. e108644, 2014.

[21] M. Sommer, A. B. Caruso, M. Sahin, T. Calderon, S. Cavill, T. Mahon and P. A. Phillips-Howard. "A time for global action: Addressing girls' menstrual hygiene management needs in schools". *PLoS Medicine*, vol. 13, no. 2, p. e1001962, 2016.

[22] A. M. Van Eijk, M. Sivakami, M. B. Thakkar, A. Bauman, K. F. Laserson, S. Coates and P. A. Phillips-Howard. "Menstrual hygiene management among adolescent girls in India: A systematic review and meta-analysis". *BMJ Open*, vol. 6, no. 3, p. e010290, 2016.

[23] UNICEF. "*Guidance on Menstrual Health and Hygiene"*. Sec. 2. UNICEF; 2019. Available from: https://www.unicef.org/media/91341/file/unicef-guidance-menstrual-health-hygiene-2019.pdf [Last accessed on 2025 May 03].

ORIGINAL RESEARCH ARTICLE UHD JOURNAL OF SCIENCE AND TECHNOLOGY

Enhancing Oat Yield and Yield components by Salicylic Acid in Different agro-ecosystem

Arol M. Anwar*

Department of Biology, Faculty of Science and Health, Koya University, Koya KOY45, Kurdistan Region – F.R. Iraq

ABSTRACT

Climate elements including temperature, humidity, and precipitation all are effects on crop growth and development especially in the arid and semi-arid areas of Northern parts of Iraq. Recently, the stresses of global climate change appear as an effective challenge, so, this study was carried out at two different locations in the Northern part of Iraq; Koya/Erbil and Altun Kopru/Kirkuk to study the effects of salicylic acid (SA) (0, 100, 200, and 300 ppm) twice as foliar spraying on five oat varieties. Results indicate that the Koya district environment significantly improved most studied characteristics, except for panicle number, compared to Altun Kopru. Except the harvest index (HI) which was non-significant, 100 and 200 ppm of SA improved significantly all studied characteristics compared to the control and 300 ppm of SA. Kangaroo, ICARDA Short, and ICARDA Tall varieties were the tallest plants (85.969, 83.469, and 82.833 cm) respectively. Each of Possum and ICARDA Short varieties were superior in panicle number, seed yield, biological yield, and straw yield. Kangaroo variety has the lowest harvest index (42.695%)significantly compared to all other varieties, whereas ICARDA Short variety was the lowest in panicle grain weight(0.898 g/panicle^{-1}).

Index Terms: *Avena sativa* L., Arid and Semiarid Area, Climate Change, Eco-Physiology, Salicylic Acid, Phytohormones.

1. INTRODUCTION

Oat (*Avena sativa* L.),is a small but important grain cereal and forage crop, ranked sixth in grain production after maize, rice, wheat, barley, and sorghum [1]. However, it is consider as one of the important grain crops, which its importance is coming from its multiple uses, where it is consumed by human and as feed for livestock, where it is rich in starch, glucan, protein, antioxidants, and fat [2]. Iraq, especially the Northern part is considered as semiarid area [3] and it depends on rainfall for irrigation which is the main factor for decreasing grain production in the Northern part of Iraq, so, any fluctuations in the ecological characteristics especially temperature and misdistribution of precipitation affects grain production. More warmer climate shortens crop growth including the grain filling period and decreases the photosynthesize accumulation, enhancing leaves senescence and increase plant respiration [4], [5]. For each Celsius degree rise in mean air temperature global is linked to an average drop of at least 3.1% in global yields of major crops [6].

Consequently, to decrease the adverse effects of stress on plants, many strategies were applied including genetically modified plants or varieties, in addition to the use of salicylic acid (SA) to increase yield and yield component [7]–[9]. SA is a safe phenolic compound for human health and it is an endogenous plant hormone it is produced normally in very small quantities. SA regulates many biochemical and physiological processes in plants such as seed germination,

Access this article online

DOI:10.21928/uhdjst.v9n1y2025.pp114-122

E-ISSN: 2521-4217
P-ISSN: 2521-4209

Corresponding author's e-mail: Arol M. Anwar, Department of Biology, Faculty of Science and Health, Koya University, Koya KOY45, Kurdistan Region – F.R. Iraq, E-mail: arol.anwar@koyauniversity.org

Received: 02-02-2025 Accepted: 13-04-2025 Published: 06-05-2025

growth of plants, flower induction, uptake of nutrients, membrane permeability, transport, plant-water relations, stomatal conduction, photosynthesis, and enzyme activities [10]. Several researches have reported that foliar application of SA results in a significant effect on plant adaptation to stress factors, as SA actions as a natural signaling particle targeting different plant species to increase their tolerance to abiotic pressures, and in plant response to environmental stress [11], [12].

High oat crop production due to good varieties, intensive cultural management practices, and relatively good weather conditions [13]–[17].

While previous studies have explored the effects of SA on crop yields, few have focused on its application in arid and semi-arid regions, particularly for oat varieties, so the aim of this study is an attempt to investigate the effects of two different ecological conditions on yield and yield components of five varieties of oat under SA application and the possibility of exploiting low-rainfall lands for grain production and providing cheap food for humans and animals. We hypothesized that SA application would significantly improve oat yield components, particularly under the stress conditions of arid and semi-arid environments.

2. MATERIALS AND METHODS

2.1. Study Area

The study area is demonstrated in the Fig. 1 a which shows the locations of each of Koya in Erbil governorate, Iraq (latitude 36.1° N and Longitude 44.6°, Elevation 560 msal), and Altun Kopru in Kirkuk governorate Iraq (latitude 35.8° N and Longitude 44.2°, Elevation 285 msal), while Fig. 1b provides a detailed map of the study sites.

2.2. Plant Materials, Cultivation, and Treatments

A Field experiment was carried out in the winter season of 2019–2020 at two different locations; the fields of the agricultural research station at Koya/Erbil (latitude 36.1° N and Longitude 44.6°, Elevation 560 msal), and a private field at Altun Kopru/Kirkuk (latitude 35.8° N and Longitude 44.2°, 285 msal) to study the effects of four concentrations of SA (0, 100, 200, and 300 ppm) mentioned as SA0, SA100, SA200, and SA300, these concentrations of SA were selected based on previous studies showing optimal yield improvements at these levels ,where one gram of SA was added to one liter of water and mix well to get the concentration 1000 ppm (stock solution). One part of the stock solution added to nine parts of water to get 100 ppm, we take two part of stock solution add to eight parts of water to get 200 ppm, and so on three part of the stock solution added to seven parts of water to get 300 ppm. The solutions applied twice as foliar spraying until a complete wetting, first spray was applied at tillering stage, and the second at flowering stage, 0.1% detergent was added to the treatment solutions as a surfactant chemical. Five oat varieties obtained from College of Agriculture, Tikrit University were used in the study, three Australian varieties; Kangaroo, Mitika, Possums, and two varieties from ICARDA, which were ICARDA Short and ICARDA Tall.

The experimental design was a randomized complete block in split-split plots. Main plots consisted of two locations (Altun Kopru and Koya), subplots consisted of SA concentration and sub-subplots consisted of varieties; the soil was plowed and smoothed then, the planting was carried out on lines with a distance between one line and another is 20 cm. Crop services including fertilization and weed control were carried out as needed. The seeds were cultivated at 15 and 17 December 2019 in both locations, respectively, and the plants were harvested on the 20 and 22 of June 2020 after reaching the stage of full maturity in both locations, respectively.

2.3. Studied Characteristics

During the physiological maturity plants height was measured from the soil surface to the plant panicle base to calculate the average height in cm for twenty plants, and the number of panicles per square meter of midline from each experimental unit where measured by calculating the total effective spikes [18].

At the day of harvesting, yield and yield components were evaluated as it mentioned by Al-Jobouri [19] which included; panicle grain yield (g/panicle^{-1}) which was measured from one-meter length of one midline from each experimental unit, grain yield (ton/hectare) was calculated from the yield of the harvested area of one-meter length of midline from each experimental unit and converted to ton/hectare, biological yield (ton/hectare) was calculated from the weight of the entire dry plants harvested above soil and for an area of 1 m length of midline and then converted to ton/hectare. Straw yield is the biological yield without grain yield. The percent of the harvest index (HI) denotes the ratio of economic yield to biological yield and It was calculated from the following equation:

Harvest index = (grain yield/Biological yield) × 100 [20]

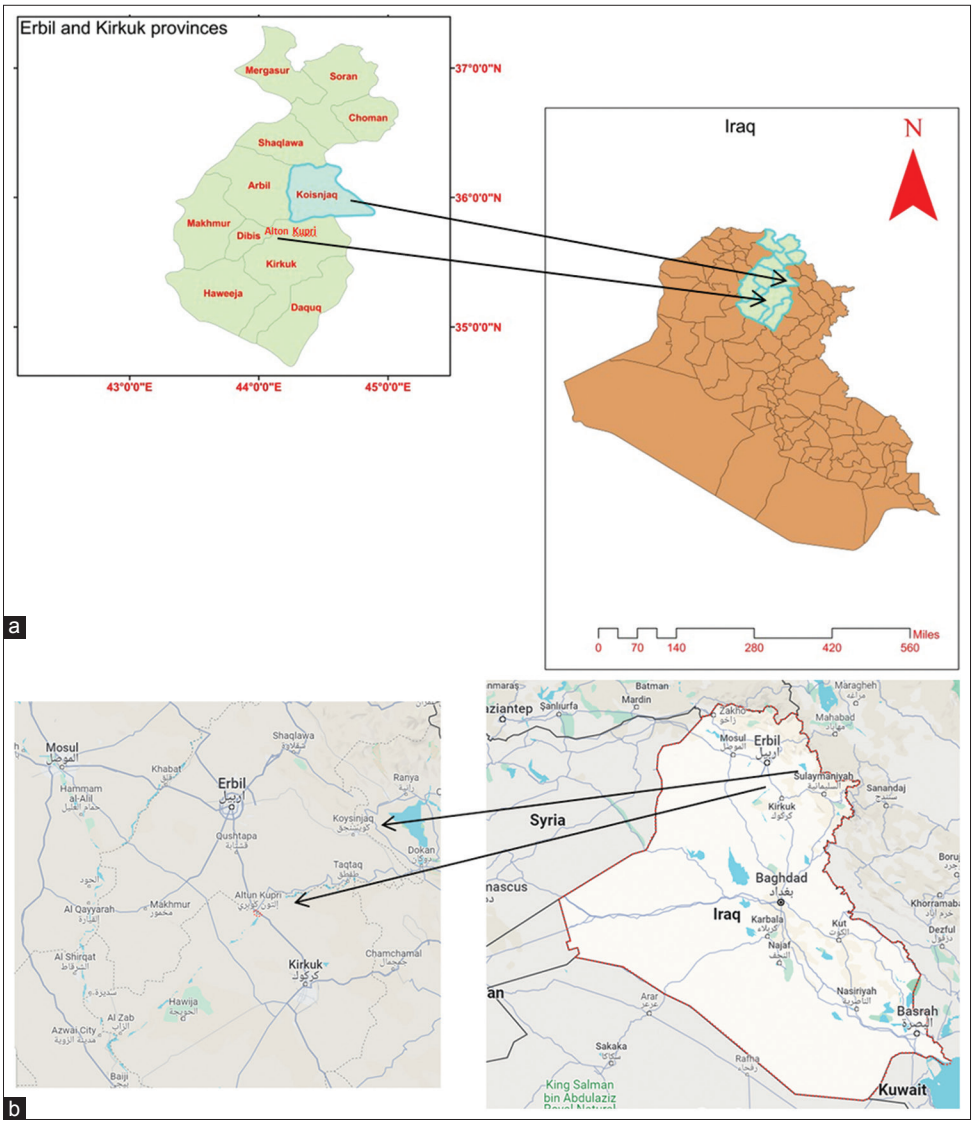


Fig. 1. (a) The areas of the study, (b) detailed map of the study sites.

2.4. Soil Characteristics and Meteorological Data

Table 1 shows some physical and chemical characteristics of the study area soils, whereas Fig. 2 shows some meteorological data of the study locations during the study period.

Soil analysis was done in the Kirkuk Agricultural Directorate in Soil Department Laboratory

2.5. Statistical Analysis and Experimental Design

Data were submitted for analysis of variance, the mean separation among treatment means for locations, SA concentration, and varieties by using Duncan's Multiple Range test at a probability level of %5 was used for comparing between experiment means using the SAS program version 9.4 [21].

3. RESULTS AND DISCUSSION

Results revealed that locations and their ecological contributions significantly affect on all studied characteristics except the straw yield. Koya district environment condition was more favorable for the plant growth and development excluding panicle number, compared to Altun Kopru district.

Regarding plant height, plant growth was better and plant height increased significantly in Koya district (74.898 cm)

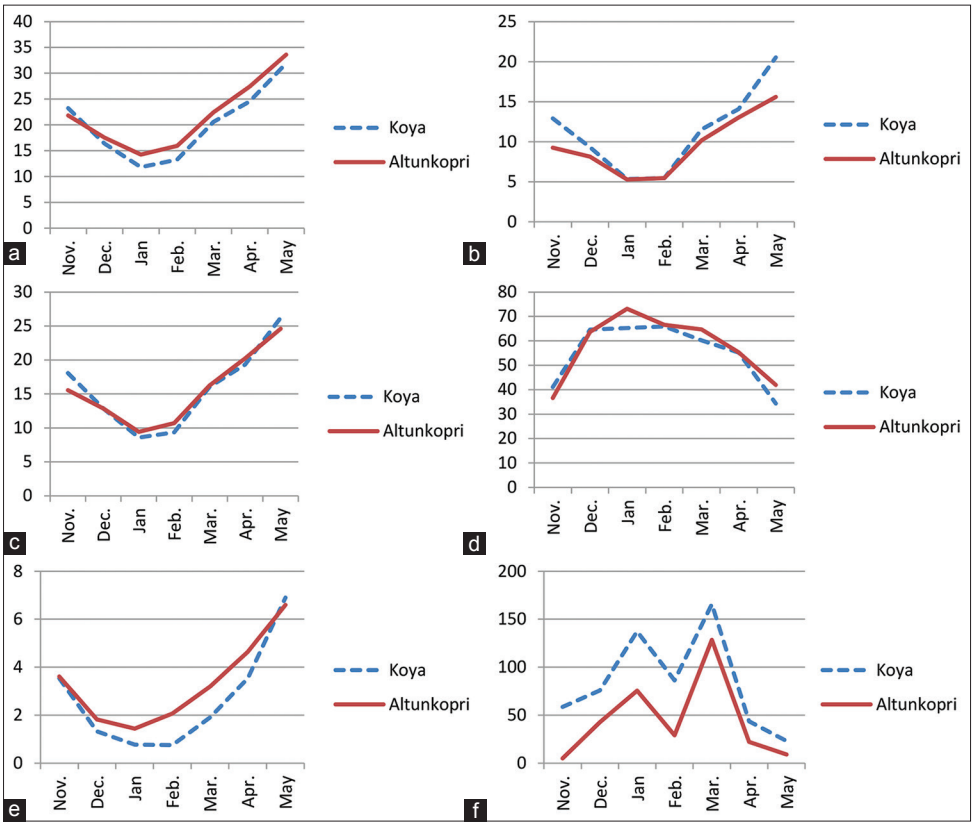


Fig. 2. Meteorological data of the study locations; Koya- Erbil, and Altun Kopru- Kirkuk During November 2019–May 2020. (a) Maximum temperature (◦C), (b) minimum temperature (◦C), (c) temperature average (◦C), (d) relative humidity (%), (e) evaporation (mm), (f) Precipitation (mm).

TABLE 1: Physical and chemical analysis for the study soils during winter season

Physical soil properties	pH	Ec (ms.cm^{-1})	O. M (%)	N (ppm)	P (ppm)	K (ppm)	Sand (%)	Silt (%)	Clay (%)	Texture
Koya	7.6	0.30	1.20	1200	8.80	120	18.3	41.1	40.6	Silty clay
Altun Kopru	6.8	0.12	0.62	350	1.26	106	66.0	16.0	18.0	Sandy loam

compared to Altun Kopru (70.646 cm)as it shown in table 2. The variety Kangaroo(85.969 cm) with ICARDA Short (83.469 cm) and Tall (82.833 cm)varieties gave highest plants significantly compared to other two Australian varieties Mitika(53.531 cm) , Possum (58.058 cm). Salicylic acid application had a significant effect in increasing plant height especially the low concentration 100 ppm which reached (75.522 cm), whereas increasing the concentration to 200 and 300 ppm which reached (73.800 ,70.875 cm) respectively led to non-significant effects compared to the control treatment (70.892 cm).Kangaroo variety regardless SA concentration gave the highest plants, and to some extend plants of each of ICARDA Short, ICARDA Tall compared to Mitika and Possum varieties.

In this study, it was shown that the number of branches was equal to the panicle number for all experimental units, so it was excluded from the result and we depend on panicle number only. Table 3 shows that the panicle number for Altun Kopru plants increased significantly to 85.033 panicle/m^{-1} compared to Koya district (79.45 panicle/m^{-1}). The Australian variety Possum, with ICARDA Short, gave the highest panicle number reached 94.313 and 90.396 panicle/m^{-1} significantly compared to all other varieties, whereas the lowest number recorded for the Kangaroo variety (68.646 panicle/m^{-1}). The application of SA regardless the concentration increased panicle number significantly compared to the control treatment which recorded the lowest value (76.367 panicle/m^{-1}), whereas the SA at 200 ppm increased them to 86.55 panicle/m^{-1} significantly compared to the concentration of 300 ppm SA.

Panicle number in the variety Possum increased significantly to 104.25 panicle/m^{-1} where they received SA at 200 and 300 ppm, whereas the variety Kangaroo that didn`t receive SA gave the lowest panicle number (59.583 panicle/m^{-1}).

As it reported in Table 3 the panicle number in the Koya district was less than Altun Kopru, which led to distribute the photosynthetase to the lowest number of panicles, as a result to increase the weight of panicles grain weight of plants grown in Koya district to 1.174 g/panicle^{-1} compared to ALtun Kopru (0.857 g/panicle^{-1}) (this effect was followed by other yield components), this effect was continue for the varsities, Possum and ICARDA Short where they gave the highest panicle number and the lowest panicle yield (Table 4) whereas the highest panicle grain yield recorded for other varieties, especially the Kangaroo

TABLE 2: Effect of SA concentrations, oat varieties, and their interactions on plant height (cm) at Koya and Altun Kopru

Varieties	SA (ppm)				Varieties means
	0	100	200	300	
Kangaroo	83.500^{a-d*}	87.500^{a,b}	88.500^{a}	84.375^{a-c}	85.969^{a}
Mitika	50.875^{g}	55.625^{f,g}	52.375^{f,g}	55.250^{f,g}	53.531^{c}
Possum	54.375^{f,g}	62.608^{e}	59.000^{e,f}	56.250^{e-g}	58.058^{b}
ICARDA short	84.750^{a-c}	83.250^{a-d}	89.125^{a}	76.750^{d}	83.469^{a}
ICARDA tall	80.958^{b-d}	88.625^{a}	80.000^{c,d}	81.750^{a-d}	82.833^{a}
SA Mean	70.892^{b}	75.522^{a}	73.800^{a,b}	70.875^{b}	
Locations mean					
Koya				Altun Kopru	
74.898^{a}				70.646^{b}	

*Means followed by the same letters for factors and their interactions are not significantly different at $P \leq 0.05$ according to Duncan's multiple range test and vice versa. SA: Salicylic acid.

TABLE 3: Effect of SA concentrations, oat varieties, and their interactions on panicles number (panicle/m^{-1}) at Koya and Altun Kopru

Varieties	SA (ppm)				Varieties means
	0	100	200	300	
Kangaroo	59.583^{j*}	77.500^{e-h}	65.000^{I,j}	72.500^{f-i}	68.646^{d}
Mitika	66.500^{h-j}	78.750^{d-g}	83.500^{c-f}	70.750^{g-j}	74.875^{c}
Possum	82.500^{c-g}	86.250^{c-e}	104.250^{a}	104.250^{a}	94.313^{a}
ICARDA short	90.083^{b-d}	99.750^{a,b}	94.500^{a-c}	77.250^{e-h}	90.396^{a}
ICARDA tall	83.167^{c-f}	80.500^{d-g}	85.500^{c-e}	82.750^{c-g}	82.979^{b}
SA Mean	76.367^{c}	84.550^{a,b}	86.550^{a}	81.500^{b}	
Locations mean					
Koya				Altun Kopru	
79.450^{b}				85.033^{a}	

*Means followed by same letters for factors and their interactions are not significantly different at P≤0.05 according to Duncan's multiple range test and vice versa. SA: Salicylic acid

TABLE 4: Effect of SA concentrations, oat varieties and their interactions on panicle grain weight (g/panicle^{-1}) at Koya and Altun Kopru

Varieties	SA (ppm)				Varieties means
	0	100	200	300	
Kangaroo	1.066^{a-e*}	1.015^{a-e}	1.176^{a-c}	1.062^{a-e}	1.080^{a}
Mitika	1.129^{a-d}	0.939^{b-e}	1.207^{a,b}	1.022^{a-e}	1.074^{a}
Possum	0.895^{c-e}	1.096^{a-e}	0.928^{b-e}	0.925^{b-e}	0.961^{a,b}
ICARDA short	0.976^{a-e}	0.975^{a-e}	0.844$^{d,\ e}$	0.798^{e}	0.898^{b}
ICARDA tall	1.129^{a-d}	1.247^{a}	1.045^{a-e}	0.826^{d,e}	1.062^{a}
SA mean	1.039^{a,b}	1.054^{a}	1.040^{a,b}	0.926^{b}	
Locations mean					
Koya				Altun Kopru	
1.174^{a}				0.857^{b}	

*Means followed by same letters for factors and their interactions are not significantly different at $P \leq 0.05$ according to Duncan's multiple range test and vice versa. SA: Salicylic acid

variety (1.08 g/panicle^{-1}). Increasing SA increased the panicle yield but non-significantly compared to the control treatment, whereas increasing the concentration to 300 ppm decreased it to 0.926 g/panicle^{-1} significantly compared to 100 ppm treatment (1.054 g/panicle^{-1}).

Each of seed yield, biological yield, and HI increased significantly in the Koya district compared to Altun Kopru, where environmental differences between them such as maximum temperature during the study period was less compared to Altun Kopru (Fig. 1), in addition, the minimum temperature in the beginning of crop establishment branching stages was higher in Koya district compared to Altun Kopru, although the average temperature was approximately similar for both districts (Fig. 1), where this different, especially in the plants' establishment was clear in the different in panicles number (Table 3).

Increasing precipitation amount and decreasing evaporation in the Koya district might be in the interest of increasing the yield and its components compared to the Altun Kopru environment. The great effects of thermal factor and precipitation on the growth and development of oat crop was also studied by [22], [23] also found that climate change in the semi-arid region leads to a negative effect on oats growth, and a reduction in oats production. Our findings are consistent with those of [24], who reported that SA application improved plant tolerance to abiotic stress.

Regarding the variety, although of their significantly low panicle seed weight (Table 4) the varieties Possum and ICARDA Short record the highest values of different yield components: Seed yield (ton/hectare), biological yield (ton/hectare), and straw yield, this increment may due to the increasing of panicle number in these varieties compared to others, which reflects on yield components. The straw yield was reflected by plant high, where the Mitika variety records significant low plant high and significant low straw yield compared to other treatments. The superior performance of Possum and ICARDA Short may be attributed to their genetic tolerance to drought and heat stress.

Increased HI has been one of the principal factors contributing to genetic yield improvements in oat (*A. sativa* L.). Although high HI demonstrates high-yielding ability when cultivars are compared, it can also indicate challenges to yield formation when comparisons are made across differing growing conditions. From this study it is clear that yield component increased in that treatments with high harvest index (Table 7) regarding the locations ecosystems, whereas, it doesn't regarding the genotypes (varieties), which is not agrees with [25] results of their study which carried out in southern Finland and reported that HI was associated with short plant stature in the studied cereal species (wheat oat, barley), and grain yield was associated closely with HI for 10 oat cultivars. However, high HI didn't indicate the degree of success in yield determination when environments are compared, which reflects the difference in environmental conditions between Finland and the Northern part of Iraq.

Most characteristics show a superior effect of SA at low (100 ppm) and moderate (200 ppm) concentrations in increasing plants growth and yield components (Tables 2-6 and 8), where in latest decades, many studies have reported that using SA as the foliar application had a significant effect on plant adaptation to stress factors, as SA actions as a natural signaling particle targeting different plant species to increase their tolerance to biotic and abiotic pressures [10]. The use of SA at 100-200 ppm could be recommended for oat cultivation in similar environments to mitigate the effects of climate stress.

TABLE 5: Effect of SA concentrations, oat varieties and their interactions on grain yield (ton/ha) at Koya and Altun Kopru

Varieties	SA (ppm)				Varieties means
	0	100	200	300	
Kangaroo	2.911^{g*}	3.638^{d-g}	3.762^{d-g}	3.292$^{f, g}$	3.400^{c}
Mitika	2.877^{g}	3.899^{c-g}	3.465^{e-g}	2.817^{g}	3.265^{c}
Possum	4.959^{a-d}	5.330^{a,b}	5.493^{a}	4.317^{a-f}	5.025^{a}
ICARDA short	4.764^{a-e}	5.097^{a-c}	5.558^{a}	4.137^{b-g}	4.889^{a}
ICARDA tall	3.709^{d-g}	4.444^{a-f}	4.043^{b-g}	3.799^{c-g}	3.999^{b}
SA mean	3.844^{b}	4.481^{a}	4.464^{a}	3.672^{b}	
Locations Mean					
Koya				Altun Kopru	
4.571^{a}				3.659^{b}	

*Means followed by same letters for factors and their interactions are not significantly different at $P \leq 0.05$ according to Duncan's multiple range test and vice versa. SA: Salicylic acid

TABLE 6: Effect of SA concentrations, oat varieties and their interactions on biological yield (ton/ha) at Koya and Altun Kopru

Varieties	SA (ppm)				Varieties means
	0	100	200	300	
Kangaroo	6.667^{f-i}*	8.938^{a-g}	8.813^{a-g}	7.688^{c-g}	8.026^{b}
Mitika	5.188$^{i, h}$	7.438^{c-h}	6.625^{g-i}	5.125^{i}	6.094^{c}
Possum	9.188^{a-e}	9.813^{a-c}	10.25^{a,b}	8.188^{b-g}	9.360^{a}
ICARDA short	9.063^{a-f}	9.521^{a-d}	10.813^{a}	7.875^{b-g}	9.318^{a}
ICARDA tall	7.084^{e-i}	8.875^{a-g}	7.688^{c-g}	7.313^{d-i}	7.740^{b}
SA mean	7.438^{b}	8.917^{a}	8.838^{a}	7.234^{b}	
Locations mean					
Koya			Altun Kopru		
8.434^{a}			7.782^{b}		

*Means followed by same letters for factors and their interactions are not significantly different at $P \leq 0.05$ according to Duncan's multiple range test and vice versa. SA: Salicylic acid

TABLE 7: Effect of SA concentrations, oat varieties and their interactions on harvest index (%) at Koya and Altun Kopru

Varieties	SA (ppm)				Varieties means
	0	100	200	300	
Kangaroo	44.072^{b}*	41.264^{b}	43.047^{b}	42.395^{b}	42.695^{b}
Mitika	55.240^{a}	52.443^{a}	52.410^{a}	54.827^{a}	53.730^{a}
Possum	51.312^{a}	54.571^{a}	53.297^{a}	52.740^{a}	52.980^{a}
ICARDA short	52.165^{a}	54.883^{a}	51.599^{a}	53.860^{a}	53.127^{a}
ICARDA tall	52.070^{a}	49.637^{a}	51.664^{a}	51.456^{a}	51.206^{a}
SA Mean	50.972^{a}	50.559^{a}	50.403^{a}	51.056^{a}	
Locations mean					
Koya				Altun Kopru	
54.015^{a}				47.480^{b}	

*Means followed by same letters for factors and their interactions are not significantly different at $P \leq 0.05$ according to Duncan's multiple range test and vice versa. SA: Salicylic acid

TABLE 8: Effect of SA concentrations, oat varieties and their interactions on straw yield (ton/ha) at Koya and Altun Kopru

Varieties	SA (ppm)				Varieties means
	0	100	200	300	
Kangaroo	3.756^{c-e}*	5.300^{a}	5.052^{a,b}	4.396^{a-e}	4.626^{a}
Mitika	2.312^{f}	3.539^{c-e}	3.160^{e,f}	2.309^{f}	2.830^{c}
Possum	4.229^{a-e}	4.483^{a-d}	4.758^{a-c}	3.872^{b-e}	4.335^{a}
ICARDA short	4.299^{a-e}	4.425^{a-d}	5.255^{a}	3.739^{c-e}	4.429^{a}
ICARDA tall	3.374^{d-f}	4.432^{a-d}	3.645^{c-e}	3.514^{c-e}	3.741^{b}
SA mean	3.594^{b}	4.435^{a}	4.374^{a}	3.566^{b}	
Locations mean					
Koya				Altun Kopru	
3.862^{a}				4.122^{a}	

*Means followed by same letters for factors and their interactions are not significantly different at $P \leq 0.05$ according to Duncan's multiple range test and vice versa. SA: Salicylic acid

Panicle grain weight, biological yield, and straw yield decreased by increasing SA concentration to 300 ppm, which may be due to decreasing physiological reactions, such as photosynthesis, transpiration, and respiration as a result to increasing ABA hormone, [26] which leads to decreasing stomata conductance, It has recently been found that SA treatment caused both ABA and proline accumulation in wheat and increased resistance to salinity [27]. Furthermore, [26] reported that SA treatment increased ABA content in the leaves of the studied oat genotypes. An increase of proline level was observed only in *Hordeum spontaneum*. The obtained results suggest that ABA and proline can contribute to the development of anti-stress reactions induced by SA. The results of this study also agree with [28] where exogenous

SA reduced transpiration and increased nitrate reductase activity, as well as the yield of some plants, also agrees with [9] where SA reduces the inhibitory effects of cadmium stress by improving the performance of root, chlorophyll a, peroxidase, catalase, proline, and total carbohydrates in lettuce (*Lactuca sativa* L.) leaves.

4. CONCLUSION AND RECOMMENDATIONS

Based on our results, we recommend the application of SA at 100-200 ppm to improve oat yields in arid and semi-arid regions, particularly for varieties such as Possum and ICARDA Short, and it reduce the inhibitory effects of the stress of ecological changed especially temperature and precipitation. Our study demonstrates that SA application significantly improves oat yield components under stress conditions, highlighting its potential as a sustainable strategy for climate-resilient agriculture. From the result of soil analysis, it appears that Koya district soil is more fertile, and all the soil parameters are more favorable for plant growth, yield, and yield components than Altun Kopru soil, which reflects clearly in the growth, and yield components.

Further studies are recommended to increase crops yield in areas suffering environmental stresses, in addition to concentrate on local and wild genotyped to achieve these goals. Also studies should explore the long-term effects of SA application on soil health and crop rotation systems in arid regions.

REFERENCES

[1] B. L. Ma, Z. Zheng and C. Ren. “Oat”. In: *Crop Physiology Case Histories for Major Crops*. Ch. 6. Academic Press, United States, pp. 222-248, 2021.

[2] M. Ahmad, G. Z. Zaffar, Z. A. Dar and M. Habib. “A review on oat (*Avena sativa* L.) as a dual-purpose crop”. *Scientific Research and Essays*, vol. 9, no. 4, pp. 52-59, 2014.

[3] M. Nanekely, M. Scholz and S. Q. Aziz. “Towards sustainable management of groundwater: A case study of semiarid area, Iraqi Kurdistan region”. *European Water*, vol. 57, pp. 451-457, 2017.

[4] J. R. L´opez, B. G. Tamang, M. Monnens, K. P. Smith and W. Sadok. “Canopy cooling traits associated with yield performance in heat-stressed oat”. *European Journal of Agronomy*, vol. 139, p. 126555, 2022.

[5] A. Raza, A. Mehmood, S. S. Razzaq, X. Zou, X. Zhang, Y. Lv and J. Xu. “Impact of climate change on crops adaptation and strategies to tackle its outcome: A review”. *Plants* (*Basel*), vol. 8, no. 2, p. 34, 2019.

[6] C. Zhao, B. Liu, S. Piao, X. Wang, D. B. Lobell, Y. Huang, M. Huang, Y. Yao, S. Bassu, P. Ciais, J. L. Durand, J. Elliott, F. Ewert, I. A. Janssens, T. Li, E. Lin, Q. Liu, P. Martre, C. Müller, S. Peng, J. Peñuelas, A. C. Ruane, D. Wallach, T. Wang, D. Wu, Z. Liu, Y. Zhu, Z. Zhu and S. Asseng. “Temperature increase reduces global yields of major crops in four independent estimates”. *Proceedings of the National Academy of Sciences of the United States of America*, vol. 114, no. 35, pp. 9326-9331, 2017.

[7] El-Taher, A.M., ABD EL-Raouf , H.S., Osman , N.A., Azoz, S.N., Omar, M.A., ELKelish , A. and ABD EL-Hady, M.A.M.(2022). Effect of salt stress and foliar application of salicylic acid on morphological, biochemical, anatomical, and productivity characteristics of cowpea (Vigna unguiculata L.) plants. Plants 11(1): 115. DOI:10.3390/plants11010115.

[8] Mujahid , M., Akram, M.Z. and Nazeer , S. (2022). Evaluation ofmung bean (Vigna radiata) under drought stress by foliar application of salicylic acid.Journal of Agriculture, Food, Environment and Animal Sciences , 3(1), pp.1-14.

[9] N. K. Talabany and I. M. Albarzinji. “Characteristics of lettuce (*Lactuca sativa* L.) under cadmium stress condition”. *Science Journal of the University of Zakho*, vol. 11, no, 1, pp. 37-44, 2023.

[10] Q. Hayat, S. Hayat, M. Irfan and A. Ahmad. “Effect of exogenous salicylic acid under changing environment: A review”. *Environmental and Experimental Botany*, vol. 68, no. 1, pp. 14-25, 2010.

[11] Shamsulleh, S.A. E.M. Al-Maaroof, and M.S. Hassan, 2016. Effect of some induce chemical and biological agents against (Tilletia tritici (Bjerk) and T.laevis (Kühn) causal agents of wheat Common bunt disease. Baghdad Science Journal, 13: 253-260.

[12] Al-Maaroof, E.M., A.H Fayadh and F.A. Fattah, 2014. Use of some chemical inducers to improve wheat resistance to Puccinia striiformis f. sp. tritici.Ratar. Povrt., 51:83-90.

[13] S. H. Al-Rubaiee. “Response of three cultivars of Oats (*Avena sativa* L.) to humic acid and its effect on yield and its components”. *International Journal of Agricultural and Statistical Sciences*, vol. 17, pp. 2201-2205, 2021.

[14] Bibi,H., Hameed,S. , Iqbal,M. , Al-BartyA., Darwish , H. , Khan,A. , Anwar, S., Mian,I.A. ,Ali,M. , Zia,A. , Irfan,M., Mussarat,M,2021. Evaluation of exotic Oat (Avena sativa L.) varieties for forage and grain yield in response to different levels of nitrogen and phosphorous. PeerJ. Doi: 10.7717/peerj.12112.

[15] ErbaŞ Kose ,Ö. D., 2022.Multi-Environment Analysis of Grain Yield and Quality Traits in Oat (Avena sativa L.).Journal of Agricultural Sciences. 28 (2) ,PP. 278 – 286. DOI: 10.15832/ankutbd.893517.

[16] V. A. Sapega and G. S. Tursumbekova. “Interaction of genotype-environment, yield and adaptive potential of Oat varieties in conditions of subtaiga of the Northern trans-urals”. *IOP Conference Series: Earth and Environmental Science*, vol. 1045, p. 012077. Doi: 10.1088/1755-1315/1045/1/012077.

[17] N. Y. Abed and H. K. S. Al-Essawi. “Ealuation of Oat varieties under sufficient and insufficient irrigation”. *Iraqi Journal of Agricultural Sciences*, vol. 55, no. 3, pp. 1251-1258.

[18] A. M. A. Wali. “*Effect of Ecological Location and Seeding Rate on Physiological Characteristics and Yield of Oat Varieties* (*Avena sativa* L.)”. Ph.D. Thesis University of Mosul, 2014.

[19] S. A. Y. G. Al-Jobouri. “*Effect of Nitrogen Fertilization and Irrigation on Characteristics of Growth, Yield and Quality of Hay and Grains of Oat Varieties* (*Avena sativa* L.)”. Ph.D. Thesis University of Mosul, 2012.

[20] R. C. Sharam and E. L. Smith. “Selection for high and low harvest index in three winterwheat populations”. *Crop Science*, vol. 26, pp. 1147-1150.

[21] S. H. Al-Mohammadi and F. Al-Mohammadi. “*Statistics and Experimental Design*”. Dar Osama for Publishing and Distribution,

Amman, Jordan, p. 375, 2002.

[22] A. Ruja, G. Gorinoiu, K. R. Suhai, A. L. Agapie, F. Salaf and C. M. Istrate-Schiller. "The effect of climate conditions on the phonological features of the autumn oat crop". *Life Science and Sustainable Development-Journal*, vol. 3, no. 1, pp. 91-97, 2022.

[23] Q. Ma, Y. You, Y. Shen and Z. Wang. "Adjusting sowing window to mitigate climate warming effects on forage oats production on the Tibetan Plateau". *Agricultural Water Management*, vol. 293, p. 108712, 2024.

[24] R. K. Odib and E. N. Dahal. "Evaluation of the biological response of the salicylic acid on growth and yield of oat". *Plant Archives*, vol. 20, no. 1, pp. 1563-1569.

[25] P. Peltonen-Sainiop, S. Muurinen, A. Rajala and L. Jauhiainenl. "Variation in harvest index of modern spring barley, oat and wheat cultivars adapted to northern growing conditions". *Journal of Agricultural Science*, vol. 146, pp. 35-47, 2008.

[26] H. Bandurska and A. Stroi ski. "The effect of salicylic acid on barley response to water deficit". *Acta Physiologiae Plantarum*, vol. 27, no. 3B, pp. 379-386, 2005.

[27] F. M. Shakirova, A. R. Sakhabutdinova, M. V. Bezrukova, R. A. Fatkhutdinova and D. R. Fatkhutdinova. "Changes in the hormonal status of wheat seedlings induced by salicylic acid and salinity". *Plant Science*, vol. 164, pp. 317-322, 2003.

[28] I. Raskin I. "Role of salicylic acid in plants". *Annual Review of Plant Physiology and Plant Molecular Biology*, vol. 43, pp. 439-463, 1992.

ORIGINAL RESEARCH ARTICLE UHD JOURNAL OF SCIENCE AND TECHNOLOGY

Hybrid E-Recommendation System for Multi-Shop Environment[i]

Hawraz Abdalla Abubakr and Kamaran Faraj

Department of Computer Science, College of Science, University of Sulaimani, Slemani, Iraq

ABSTRACT

In the Kurdistan Regional Government, most computer shops and markets conduct their marketing offline and do not have electronic systems. Nevertheless, customers live in a digital age; they often face challenges in finding products among these markets and shops. The most common question that customers ask is which shop they should purchase from. Therefore, data from five laptop stores and ratings for markets were collected to build an integrated recommender system to help customers find products and select the best store. Our proposed system is a hybrid e-recommendation system that combines machine learning techniques to provide personalized shop and product recommendations. Methods include data collection from multiple laptop shops and dataset preparation. The system uses techniques such as hybrid/blended methods using singular value decomposition and K-nearest neighbors for collaborative filtering (CF) to recommend shops and products based on customer ratings, alongside term frequency-inverse document frequency vectorization and cosine similarity for content-based filtering. The CF's performance was evaluated using metrics like RMSE = 0.14 and MAE = 0.11, which demonstrated positive results for product and market recommendation. Overall, this study offers solutions through HE-RS to address key challenges such as market fragmentation, cold-start problems, and data scarcity.

Index Terms: Hybrid E-Recommendation System, Content-based Filtering, Collaborative Filtering, Machine Learning, E-Commerce

[i]This paper is based on the master's thesis, "Hybrid E-Recommendation System for Multi-Shop Environment," by Hawraz Abdullah Abubakr.

1. INTRODUCTION

In the age of rapid digitalization, E-recommendation systems (E-RS) have emerged as indispensable tools in most fields, from e-commerce and entertainment to education, healthcare, and restaurants. These systems aim to reduce complexity by making relevant options easily available to users or by tailoring content to personal needs [1]. As consumer choice expands exponentially, the demand for efficient, accurate, and personal recommendation systems continues to grow.

Recommendation systems between people are often very effective for mutual assistance [2]. Therefore, it can be said that the idea originated from real-world human interactions. The recommender system appeared shortly after the World Wide Web was created, and both business and academics have investigated and used related technology extensively. One of the most popular online applications nowadays, recommender systems help billions of users every day by suggesting various types of material, such as news feeds, videos, e-commerce goods, music, movies, books, games, friends, jobs. These triumphant tales have demonstrated that recommender systems are capable of transforming large amounts of data into valuable information [3].

Access this article online

DOI:10.21928/uhdjst.v9n1y2025.pp123-134

E-ISSN: 2521-4217
P-ISSN: 2521-4209

Corresponding author's e-mail: Hawraz Abdalla Abubakr, Master's Student, Department of Computer Science, College of Science, University of Sulaimani, Slemani, Iraq. Email: hawrazshawry@gmail.com

Received: 23-03-2025 Accepted: 26-04-2025 Published: 01-06-2025

Field has shifted from basic content-based filter (CBF) methods to advanced hybrid models that integrate CF, machine learning, and context-aware techniques. While CBF matches user preferences with item features, CF predicts preferences by analyzing the behavior of similar users. Despite their effectiveness, both methods face challenges such as cold-start issues and limited data, which can reduce their practical applicability.

The hybrid approach combining different filtering methods has been a significant step in overcoming these challenges. By combining insights from user behavior, product characteristics, and context, hybrid systems provide recommendations that are not only more accurate but also more diverse [4]. This is particularly helpful in regions, such as the Kurdistan Region, where fragmented markets make it difficult for consumers to find the products or services, they need across many different applications and systems.

Despite all efforts, the implementation of recommendation systems in fragmented markets is still understudied. Most present systems are designed for centralized e-commerce platforms with substantial user data, leaving a significant gap in these areas because markets operate independently and data are scarce and scattered. Addressing these gaps requires careful solutions and research that adapts to the unique challenges of fragmented ecosystems, enabling consumers to benefit from the personalized and detailed recommendations they expect in more developed markets.

2. E-RS

Recommender systems or E-RSs (sometimes rendered by terms such as platforms engines, algorithms, or software tools and techniques) are a subclass of data filtering systems that provide recommendations for items most relevant to a particular user. The aim of a recommender system is to create meaningful recommendations to a list of users for items or products that might interest them. These systems play an effective role in various areas, including e-commerce [5].

Typically, recommendations refer to various decision-making processes about what products to buy, what music to listen to, or what online news to read. RSs are used in many fields, with well-known examples taking the form of playlist generators for video and music services, product recommenders for online stores, content recommenders for social media platforms, open web content recommenders, or food recommendations in online restaurants. These systems can operate using a single type of input, such as music, or multiple inputs into and from across platforms, such as news, books, and search queries. There are also popular recommendation systems for specific topics, such as restaurants and online dating. Recommendation systems have also been developed to investigate research topics and experts, colleagues, and financial services [5].

2.1. Literature Review

Recommendation systems, as a field, have evolved rapidly, with several works addressing critical challenges, such as data sparsity, cold-start, and personalization, etc. These works are also closely related to this research, which employs a hybrid recommendation system (HRS) to improve poorly integrated and fragmented markets. In line with the overarching aim of our research, we reviewed some studies that offer insight and contribute to research goals.

Hasan and Ferdous (2024) highlight the effectiveness of HRSs that combine text-to-number transformation, matrix factorization techniques such as ALS, and cosine similarity for precise recommendations. Their study focuses on preprocessing, similarity calculations, and latent factor modeling, using the TMDB 5000 dataset and RMSE measures for evaluation. Their study further demonstrates how combining multiple methods in a hybrid approach improves the efficiency and accuracy of recommendations, especially in addressing data scarcity and enhancing system robustness [6].

Mouhiha *et al.* (2024) explore the combination of CF and CBF to address their individual limitations, such as the problem of cold start and data scarcity. Their study indicates that a HRS effectively improves accuracy and user satisfaction. Also, they describe various hybrid methods that tailor recommendations to user preferences by utilizing both user-item interactions and item specificity [7].

Loukili *et al.* (2023) investigate recommendation systems designed for e-commerce, using the FP-growth algorithm to analyze purchase patterns and generate personalized recommendations. Their study describes the decision-making challenges posed by the large number of product options. The rating metrics, including accuracy and conversions generated, also demonstrate how robust recommender systems have a direct impact on increasing sales and customer attraction, and highlight practical applications in real-world e-commerce platforms [8].

Cherkaoui *et al.* (2024) discusses using machine learning to predict customer behavior and improve strategies for marketing. This is achieved through the use of Apriori algorithms to mine the association rules and more

sophisticated techniques, such as neural networks that are used to analyze user data to enhance understanding of the behavior of the users. Furthermore, their system concentrates on data segmentation, target estimation, and measures of effectiveness, thereby assisting businesses with realistic recommendations and improving the relevance of the goods recommended to users in highly competitive markets [9].

Ozturk *et al.* (2024) present a CrossGR model integrated with Graph Isomorphism Networks (GINs), facilitating the function of cross-market recommendation systems even with data sparsity and the limitations posed by the single markets. By employing GINConv layers and multi-layer perceptrons, the model derives insights into the typically complex user-item engagements, thus generating effective recommendations for different markets. Their novel concept in recommendation systems has demonstrated significant improvements in metrics, such as NDCG@10 and HR@10, suggesting the efficacy of graph-based learning approaches in recommendation systems [10].

The reviewed literature highlights various approaches, such as (hybrid recommendation models, machine learning approaches, graph-based methods, and NLP sentiment analysis). While these studies address issues such as data scarcity, personalization, and cold start issues, they do not specifically focus on fragmented market integration, a major challenge in many parts of the world, such as the Kurdistan Region.

This thesis aims to addressee the gap by combining different techniques such as CBF and CF, designed to enhance market recommendations in a complex multi-store environment. By collecting product information from multiple markets, the proposed system provides consumers with a unified shopping experience, overcoming the problems of fragmentation in the domestic e-commerce sector. Customers can also access multiple stores and products through a single application, which reduces customer search time and storage, especially on smartphones.

2.1.1. CBF

A content-based recommendation system [11] is a type of RS that functions based on the content and characteristics of items. For instance, each product has a set of attributes, and the system determines which attributes have contributed to the user satisfaction in the past.

2.1.2. Collaborative filter (CF)

CF is a powerful tool used to personalize web experiences based on other users' preferences and opinions. This method, known as CF, helps in examining large amounts of data by combining the perspectives of vast and interconnected online communities to make recommendations [12].

CF is divided into model-based and memory-based. Each has its own style and set of challenges. Memory-based systems use past evaluations to find users with similar tastes, while model-based systems use algorithms, such as genetic algorithms, to improve the accuracy of their recommendations.

CF applications: CF is widely used across various domains to provide personalized recommendations based on user preferences and behaviors, such as E-commerce, Music and Movie Recommendations, News, social media, etc.

2.1.3. HRS

HRSs combine multiple recommendation strategies to maximize their strengths. The choice of strategy depends on data characteristics. By effectively merging methods, HRSs deliver more accurate, adaptive, and user-personalized recommendations [4]. In another way, they are blended or mixed recommender systems.

Advantages of HRS over traditional RSs:

(i) Improved accuracy and personalization: By merging several recommendation algorithms, E-HRSs can provide recommendations that are more accurate and personalized based on the user's interests and preferences, surpassing the limitations of a single RS.
(ii) Scalable and flexible: HRS is more scalable than RS alone, as it can support more users and is also capable of being flexible by providing recommendations based on changes in customer behavior and application.
(iii) Enhanced robustness: As previously mentioned, the primary issues with a single RS that prevent it from making precise suggestions are data sparsity, biases, and inconsistencies. In contrast, HRS can operate well in this situation and more effectively handle modifications to user preferences, behavior, item attributes, and other elements.
(iv) Greater coverage: While single recommendation systems are constrained in their recommendation area and are only able to suggest certain goods, HRSs are able to go beyond these limitations and offer recommendations that are wider and more comprehensive.

Due to these benefits, researchers and businesses continue to refine HRSs, focusing on accuracy, algorithm combinations, and adaptability [4].

2.1.4. Machine learning

Machine learning (ML) is a subpart of artificial intelligence (AI) that focuses on developing algorithms and models that enable computers to learn from data and make predictions or decisions without being explicitly programmed [13]. ML has revolutionized the field of E-recommendation systems, making the systems more intelligent, more accurate in personal recommendations and enabling them to understand changes faster. E-commerce has expanded significantly due to its benefits, which has created a very good environment for the use of machine learning as a branch of AI. Due to the expansion of available information, there is a need for personalized experiences, addressing fraud detection and security challenges, enhancing supply chain optimization capabilities, and recognizing the importance of customer sentiment analysis [14]. Table I shows some ML models and algorithms that are used for E-RSs.

TABLE I: Machine learning models and algorithms used for E-RSs

Model/Algorithm	Description
K-means clustering	A popular unsupervised algorithm used to group similar users/items for recommendations.
Cosine similarity with K-NN	A similarity measure used in both content-based and collaborative filtering to compare vector closeness.
TF-IDF and cosine Similarity algorithms	The TF-IDF algorithm is used to evaluate the importance of words in a textual corpus. The cosine angle between two vectors is used by the cosine similarity method to calculate how similar they are.
Matrix factorization(MF)	Decomposes the user-item matrix into latent factors for collaborative filtering and hybrid systems.
Singular value decomposition (SVD)	A dimensionality reduction method used to improve the accuracy of collaborative filtering.
Latent Dirichlet Allocation (LDA)	A topic modeling technique used for literature or text-based content recommendation systems.
Convolutional neural networks (CNN)	Used to process structured data, such as images or sound for personalized music or movie recommendations.
Support vector machines (SVM)	A supervised learning algorithm used for classification in hybrid recommendation systems.
Gradient descent	Optimization method used in collaborative filtering to minimize error in rating prediction.
Reinforcement learning	Models' user interaction dynamically using contextual multi-armed bandit techniques for real-time recommendations.
Autoencoders	A neural network used in hybrid systems to reduce sparsity and integrate multiple recommendation approaches.
Bayesian networks	A probabilistic approach for predicting user preferences based on dependencies among variables.
Funk-SVD	A variation of SVD optimized for large-scale collaborative filtering by focusing on smaller matrices.
Decision trees	A supervised learning method used in collaborative filtering to classify user preferences.

3. METHODOLOGY AND SYSTEM ARCHITECTURE DESIGN

The system begins with data collection, followed by the design of a hybrid recommendation model aimed at producing effective results for the given problem. The methodology integrates two main approaches: CBF and CF.

3.1. Data Collection and Preparation

Figure 1 show the data collection and preparation process for five laptop shops and 1000 user ratings for shops and products. The *product datasets* include information about products from different shops. Each product dataset includes features such as (market name, brand name, processor type, processor brand, generations, RAM, storage capacity in GB, have SSD, HDD, graphics capacity, price, display size, display type, and color). The *User rating dataset* includes customer ratings. This allows the system to generate more accurate recommendations based on user preferences. It contains the user ID, shop ID, product ID, and the rating given by the consumer to the shops and products.

Figure 1 also illustrates the main preprocessing steps applied. Data preparation and preprocessing began by combining the datasets from five shops into a single dataset. Subsequently, data cleaning (such as removing duplicates, punctuation, and handling missing values), normalization (e.g., for price and rating), transformation (such as converting the rating matrix to a tabular format), and the selection of key features for building the e-recommendation system based on the features Kurdish users consider important when purchasing laptops were performed. In the final step, the rating dataset was divided into a training set (80%) and a testing set (20%) to evaluate model performance effectively.

3.2. Term Frequency-Inverse Document Frequency (TF-IDF) and Cosine Similarity

TF-IDF is a text analysis method used to calculate the relevance of words in a document when compared to a collection of documents. IDF reduces the weight of frequent words that exist in multiple documents, while TF measures

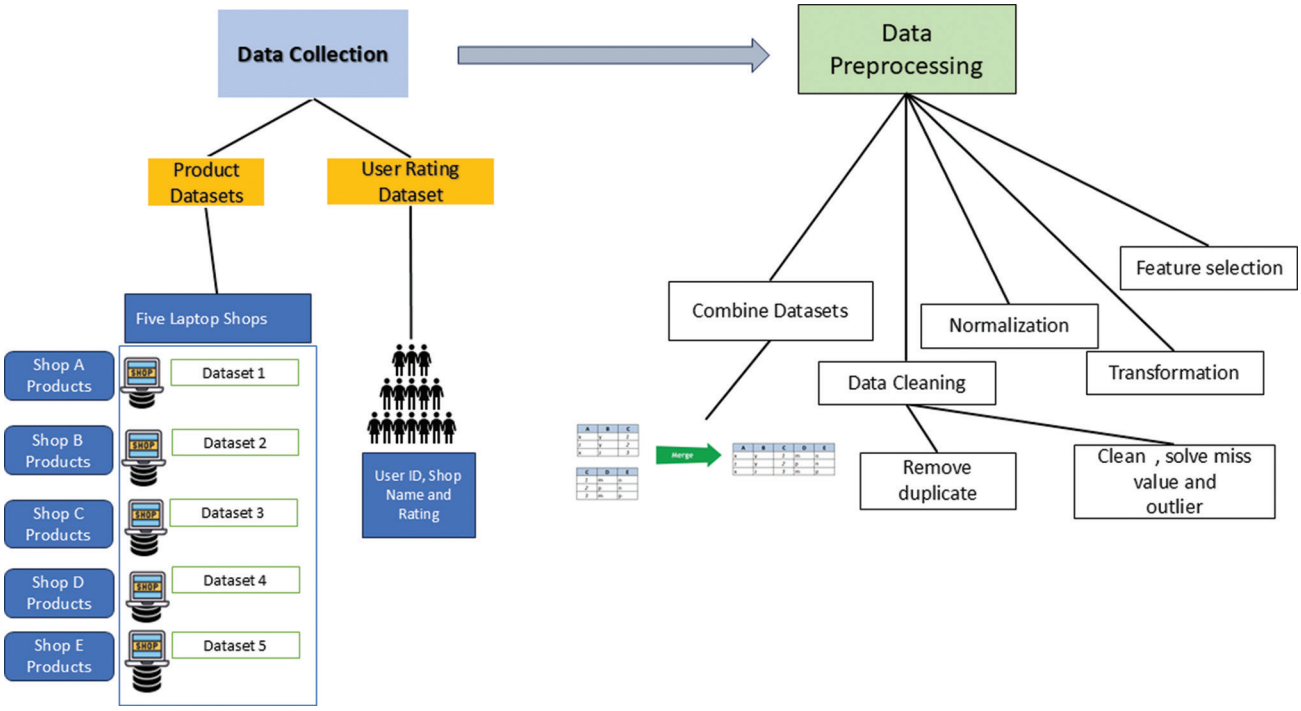


Fig. 1. Data collection and preparation process.

the frequency of a word's occurrence in a document. The TF-IDF equation is:

$$\text{TF-IDF} = \text{TF} \times \text{IDF}$$

Where is calculated as the number of times a word appears in a text or document divided by the total number of terms in that document, and IDF is computed as:

$$IDF = \log\left(\frac{N}{df}\right)$$

Where N is the total number of documents, and df is the number of documents containing that word. TF-IDF helps in filtering out less important words while highlighting meaningful ones in recommendation systems.

A measure called cosine similarity is used to evaluate how similar two text vectors are to each other as well as providing a quantitative measure of document similarity by calculating the cosine of the angle formed by two vectors in n-dimensional space. The formula is:

$$\cos(\theta) = \frac{A \cdot B}{|A||B|}$$

A and B are document vectors, and the numerator's dot product is divided by the product of their scales. Recommendation systems use both methods to detect product similarities based on descriptions in texts [15].

3.3. Singular Value Decomposition (SVD)

SVD is a matrix factorization algorithm that converts a matrix into three matrices for efficient representation and dimensionality reduction. This technique has found extensive applications across various areas, such as recommendation systems, where it has been shown to recommend new products and encourage repeat purchases, resulting in an improved user experience [16].

SVD in a recommendation system:
Input: A matrix where rows are users, columns are items (like movies), and the values are ratings.

SVD splits the matrix into three smaller matrices:
$A = U \cdot S \cdot V^T$
U: Describes the users.
S: Contains the importance of features.
V_T: Describes the items.

Using this, we can predict missing values in the original matrix, such as movies a user hasn't rated yet.

3.4. K-Nearest Neighbors (K-NN)

K-NN is an instance-based learning algorithm that is widely used in classification and regression problems [17]. Because it is so easy and supports large-dimensional data, it is widely used for text classification, image recognition, and medical diagnosis. Although algorithm performance may depend significantly on the distance used, it has been shown to achieve competitive results in most real-world applications. K-NN for recommendation system is a neat application, in which you can detect similar users or objects from their characteristics and propose them recommendations.

3.5. CBF Process

We combined multiple product attributes (such as price, brand, RAM, screen size, etc.) into a single row for each product. This row represents a general description of the product.

1. The TF-IDF Vectorizer is used to convert these combined product features into numerical vectors. This conversion transforms the textual information into a format suitable for similarity calculations. As the process is shown in Figure 2.
2. After converting product attributes and search queries into TF-IDF vectors, cosine similarity was used to compute and find the products that are most similar to the user's search query.
3. After generating initial recommendations, the results are filtered based on user-supplied parameters, such as RAM, screen size, and storage. Finally, the shops and products were sorted by their ratings and similarity score. Furthermore, Figure 2 illustrates the integration of TF-IDF and Cosine Similarity.

3.6. CF Process

The user ratings are normalized using normalization techniques to ensure all ratings are within the range of 0 to 1. This process reduces bias arising from different rating scales across users.

1. SVD is employed to analyze a sparse user-item evaluation matrix into latent factors representing user and market preferences. It enables the prediction of missing assessments by capturing underlying patterns in the data.
2. K-NN is also utilized to identify users or similar shops based on their evaluation patterns. This method offers interpretability by highlighting neighbors.
3. We combined SVD and K-NN to form a hybrid model. The SVD and K-NN predictions were weighted (40% SVD, 60% K-NN) to generate the final recommendation. This approach exploits the strengths of both algorithms to reduce errors and improve the accuracy of recommendations. Figure 3 illustrates the process flow of the CF approach, which integrates SVD with K-NN.

3.7. The Flowchart of the Proposed System

In Figure 4, the flowchart shows the user workflow in the E-RS. To begin with, the user can either create an account by

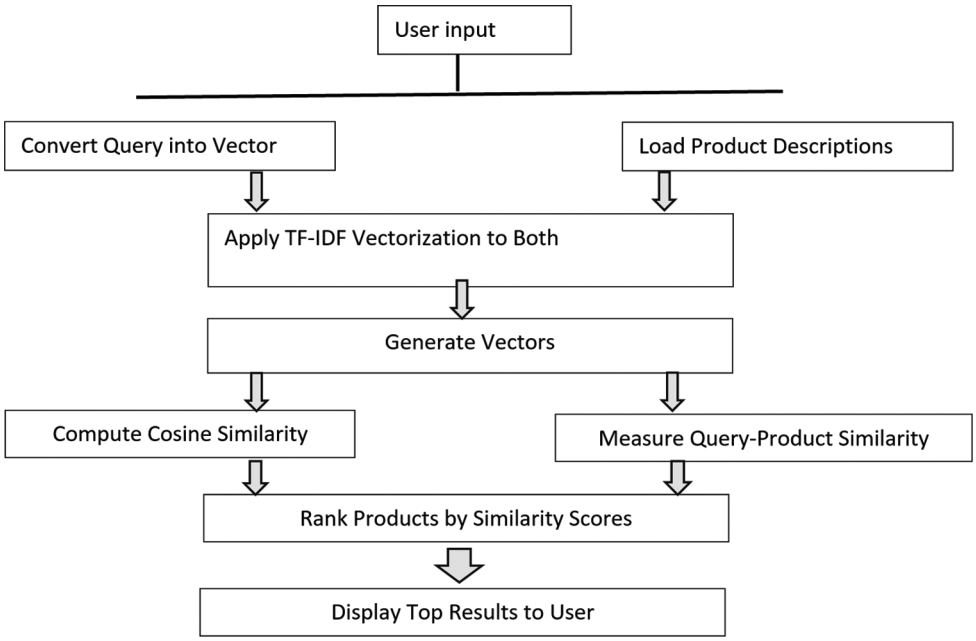


Fig. 2. Content-based filtering process.

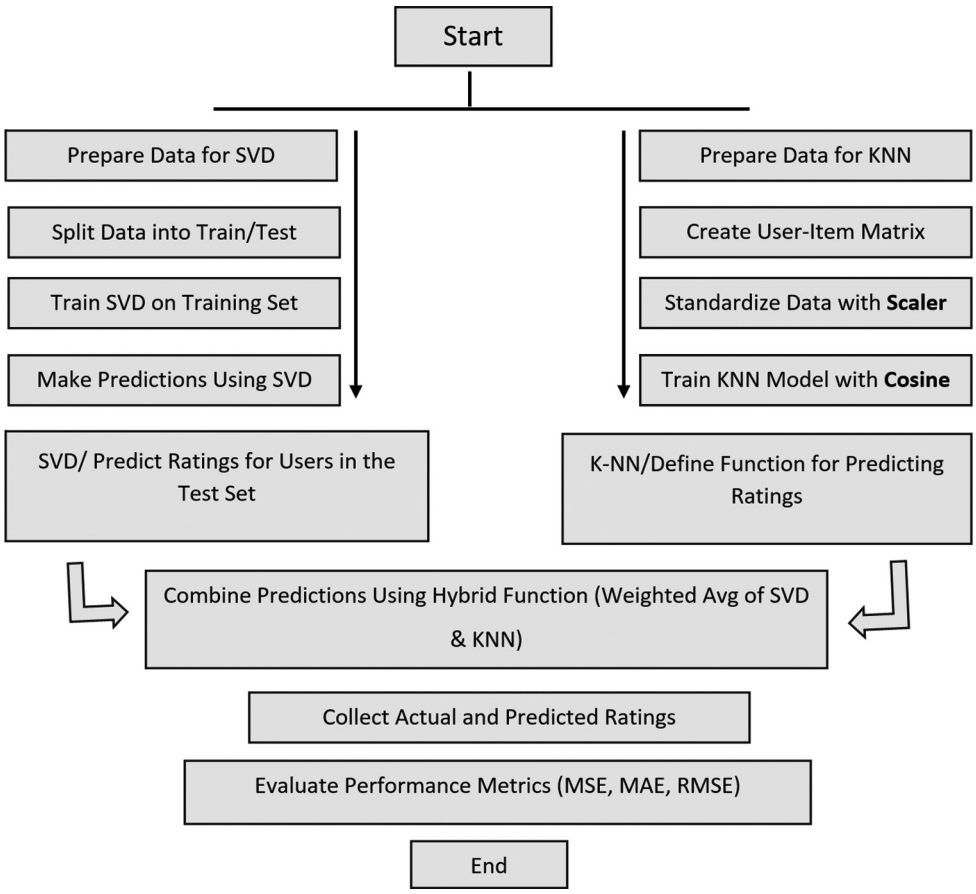


Fig. 3. Collaborative filter process.

registering or log in to their present account. Alternatively, they can search the system anonymously, limiting the level of personalization. Once a user is engaged, they can start browsing products. The system uses a combination of recommendation techniques. For example, if a user searches for "RAM 32GB", the system uses CBF to analyze user search queries and compares them with product descriptions using techniques such as TF-IDF and cosine similarity, which helps the user to find laptops in multiple shops with similar attributes such as RAM, storage, and price. The user search result is ordered based on similarity score. For example, if shop A has a high similarity score, it will be shown at the top for the user. The system also makes use of CF for registered users to analyze past user behavior, such as ratings and compares it to the behavior of other users with similar interests. Algorithms, such as K-NN and SVD are used to detect these similarities and predict the shop and products that the consumer might like.

Finally, the hybrid recommendation system combines the strengths of both CBF and CF. This approach offers a more nuanced and accurate set of recommendations by taking into account both product characteristics and user ratings.

4. RESULTS AND DISCUSSION

```
# TF-IDF Vectorizer to convert text to numerical vectors
tfidf = TfidfVectorizer(stop_words='english')

# Fit and transform the combined features for all products
tfidf_matrix = tfidf.fit_transform(products_df['combined_features'])

# Transform the search query into the same TF-IDF space
query_tfidf = tfidf.transform([search_query])

# Compute cosine similarity between the search query and all products
cosine_sim = cosine_similarity(query_tfidf, tfidf_matrix)
```

4.1. CBF Implementation and Result

Figure 5 illustrates the top five recommended Apple products retrieved using CBF. The system computed the similarity scores using TF-IDF and cosine similarity. Higher scores indicate stronger textual similarity to the user's search. This figure shows that the most relevant match (0.773) is from the

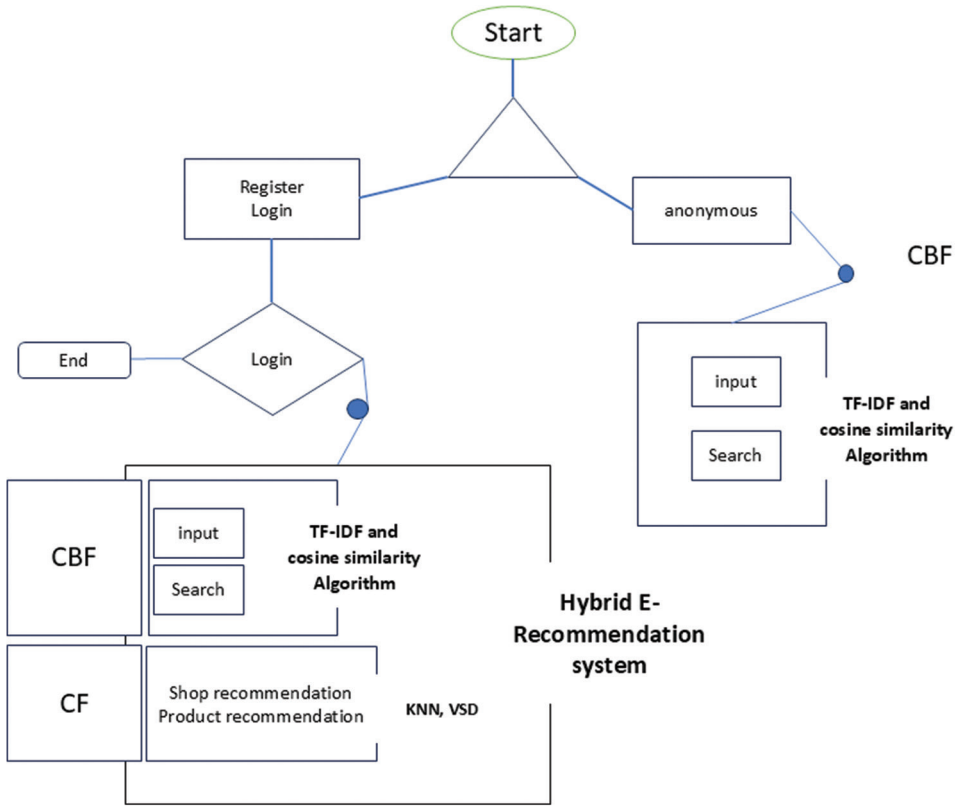


Fig. 4. Flowchart of the proposed system.

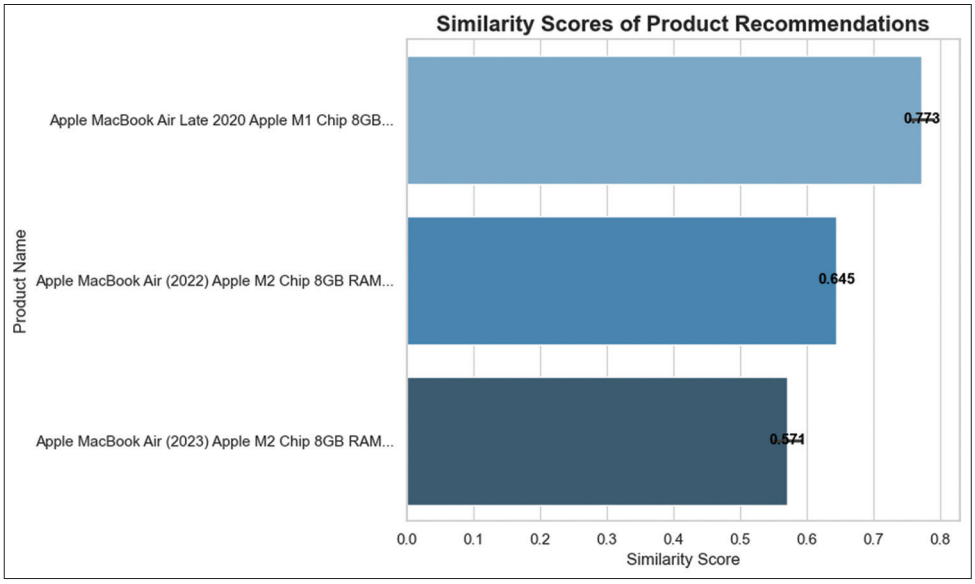


Fig. 5. Content-based filter similarity score result.

"Shop W" market, while other competitive options appear in "Shop R" and "Shop M_A".

The brief code sample in the screenshot above presents a fundamental CBF code structure. The CBF component

relied on TF-IDF to convert textual product descriptions into numerical vectors that were then used to calculate cosine similarity. Cosine similarity measures the angle between two vectors in multidimensional space which provides a measure to find the similarity between user search and product attributes. These detailed calculations ensure that the best ranking related to user search is provided.

To evaluate the CBF approach in a real-world user-facing system, we consider an example of a user search with loosely defined features ("8 GB RAM core i5 gen 11 white hp"). As observed, this query exhibits characteristics of a user search with loosely defined features. The search results for this query were available from five distinct vendors. Figure 6 presents the results from a selection of these shops.

4.2. CF Implementation and Result

The model needs to be trained and tested after it is built. The dataset was partitioned into a testing set (20%) and a training set (80%). For evaluating the models, metrics, such as root mean squared error (RMSE), mean absolute error (MAE), and mean squared error (MSE) were used. The below screenshot for the sample code shows how SVD and K-NN are combined to improve shop and product recommendations.

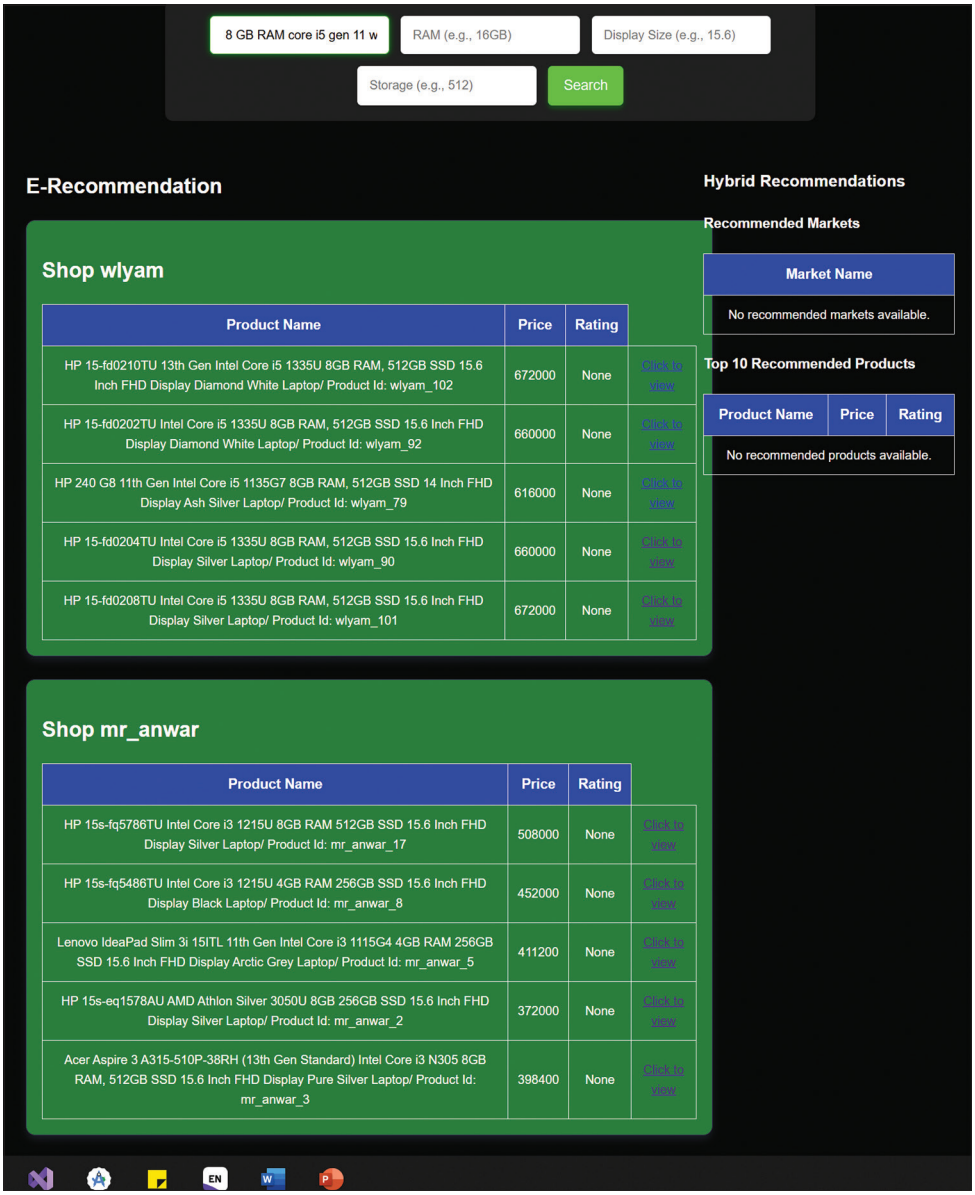


Fig. 6. User search example and content-based filter result.

```
#CF Result combination
# Combine predictions from KNN and SVD
def hybrid_predict(user, market):
    # Get SVD and KNN predictions
    svd_rating = svd_predicted_ratings.get(f"{user}-{market}", 0)
    knn_rating = knn_predict(user, market)

    # Weighted average of the two predictions
    return 0.4 * svd_rating + 0.6 * knn_rating
```

In addition to these results in table 2, the response time was also tested to evaluate the real-time performance of the hybrid proposal model. The test yielded an average prediction response time of 0.0020 s, or 2 ms, demonstrating the system's efficient and near-instantaneous shop recommendation capability. This outcome suggests that the model is effectively optimized for real-time applications, enabling it to handle user demands with minimal latency.

4.3. Hybrid E-Recommendation System Result

By bringing together both CBF and CF into one system, we create a hybrid recommendation model that takes the best of both worlds while reducing each method's weaknesses. The HE-RS model combines the feature-based approach from CBF, which uses product attributes to find similarities, and the behavior-based insights from CF, which look at user ratings and patterns to offer more personalized and robust recommendations. This means that even if a user hasn't rated many shops and products, CBF can still suggest items based on product features, while CF can step in when product details alone aren't enough to capture user preferences. Contextual and behavioral data are both included in the final recommendation outcome, which produces a system that can effectively handle problems, such as data lacking, new users, and modifying preferences. Furthermore, this hybrid approach not only improves the accuracy of recommendations but also creates a more intuitive and user-friendly shopping experience by continuously matching product recommendations and shop recommendations to changing consumer preferences. In Figure 7, what is on the right side of the system is the result of a customer's recommendation based on the rating they gave, in which machine learning algorithms play a role. The list on the left shows the results of the customer's search and the results according to the similarity of the user's search and products. In all stores, products are recommended by the recommender system based on CBF.

4.4. Discussion

Table 2 is shows the hybrid model (SVD with K-NN) gives better recommendations than the SVD model alone. Hybrid model has a lower RMSE (0.1435) and MAE (0.1147) compared to the SVD model, indicating that the hybrid approach makes more accurate predictions. SVD finds hidden patterns in the data, while K-NN improves prediction by comparing similar users. Thus, combining both methods can reduce errors and give better recommendations. The results of this study are presented in two ways that reflect real-world applicability. One of the important aspects that distinguishes this research from other studies is the development of a comprehensive system that integrates multiple recommendation techniques to enhance performance. The evaluation and results suggest that the prototype system is suitable for further development and optimization, and that a centralized recommendation platform can serve as a practical solution for mitigating challenges in fragmented markets. Instead of requiring each shop to develop its own application, which causes data dispersion, businesses can simply create an account within a unified system tailored to their specific industry, improving accessibility and user experience.

4.5. Comparative Analysis

To validate the performance of our hybrid recommendation system, we compare it with recent models that also apply hybrid or advanced recommendation techniques. As shown in Table III, our system combines SVD and K-NN for CF with TF-IDF and Cosine Similarity for CBF. The CF achieves a low RMSE of 0.14 and MAE of 0.11 on a sparse, custom laptop dataset. Furthermore, the CBF achieves successful performance in responding to customer searches. Compared to other studies, our system demonstrates higher precision and adaptability in market-specific recommendations.

TABLE II: Cross-validation result

Model	Metric	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Mean	Standard
SVD	RMSE	0.3185	0.3123	0.3184	0.3192	0.3179	0.3172	0.0025
	MAE	0.2713	0.2637	0.2741	0.2767	0.2688	0.2709	0.0045
Hybrid (SVD+K-NN)	RMSE	—	—	—	—	—	0.1435	—
	MAE	—	—	—	—	—	0.1147	—

RMSE: Root mean squared error, MAE: Mean absolute error, SVD: Singular value decomposition, K-NN: K-nearest neighbor

Fig. 7. Hybrid e-recommendation screen result.

TABLE III: Comparison of HE-RS with other studies

Studies	RMSE	MAE	Dataset	CF method	CBF method	Key insights
HE-RS	0.14	0.11	Laptop market (Custom, Sparse)	SVD, K-NN	TF-IDF, Cosine similarity	Accurate hybrid recommendations tailored to sparse, real-world e-commerce
Hasan and Ferdous [6]	0.8951	–	TMDB 5000	ALS	Cosine similarity	Hybrid approach; improved accuracy compared to basic CF
Mouhiha *et al*. [7]	0.22 (RMS Prop), 0.14 (SGD)	–	Movie lens	–	Deep features, cosine	Deep learning improved hit rate (0.89), NDCG@10=0.64
Ozturk *et al*. [10]	–	–	Cross-market dataset	Graph isomorphism networks	–	Best NDCG@10=0.6524, HR@10=0.7609 in cross-market scenario

HE-RS: Hybrid electronic recommendation system, RMSE: Root mean squared error, MAE: Mean absolute error, SVD: Singular value decomposition, K-NN: K-nearest neighbor, CBF: Content-based filter, CF: Collaborative filtering

5. CONCLUSION

In this study, we have designed a hybrid electronic recommendation system (HE-RS) for a multi-shop environment, such as the Iraq-Kurdistan region, to address market fragmentation. We have combined methods within the system, such as CBF and CF. The system allows users to find shops and products based on shop ratings and product characteristics. The results indicate that the system can simplify customer search, and algorithms such as TF-IDF and cosine similarity were able to produce accurate results for customer queries across multiple shops. The system also improved the quality of recommendations by using advanced machine learning techniques, such as the combination of SVD and K-NN algorithms to analyzing customer reviews, which effectively predicted customer preferences and achieved high performance. Moreover, the system was able to address challenges such as the cold start problem, limited data, and diverse user preferences. A major solution to using multiple applications for various shops is provided by this system. Stores can register on a centralized platform, which allows customers to browse for items in many marketplaces using a single interface; this reduces the customer search time.

5.1. Future Work

a. Integrate real-time user feedback.
b. Use deep learning techniques to improve recommendation accuracy.
c. Analyze more user behavior, such as clicks or browsing history.
d. Integrate the system with mobile applications.

Finally, our research demonstrates that a HE-RS offers a means to address market fragmentation, provides effective recommendations for customers, and generates increased opportunities for businesses.

REFERENCES

[1] F. Ricci, L. Rokach and B. Shapira. Recommender Systems: Techniques, Applications, and Challenges. In: "*Recommender Systems Handbook*". Springer Science and Business Media, Germany, pp. 1-35, 2021.

[2] S. Moica, C. Veres Harea and L. Marian. "*Effects of Suggestion System on Continuous Improvement: A Case Study*". IEEE, United States, pp. 592-596, 2018.

[3] Z. Dong, Z. Wang, J. Xu, R. Tang and J. Wen. "*A Brief History of Recommender Systems*". [arXiv Preprint], 2022.

[4] K. Bodduluri, F. Palma, A. Kurti, I. Jusufi and H. Löwenadler. "Exploring the landscape of hybrid recommendation systems in E-commerce: A systematic literature review". *IEEE Access*, vol. 12, pp. 28273-28296, 2024.

[5] F. Ricci, L. Rokach and B. Shapira, editors. Recommender Systems: Techniques, Applications, and Challenges. In: "*Recommender Systems Handbook*". New York, Springer US, pp. 1-35, 2022.

[6] M. R. Hasan and J. Ferdous. "Dominance of AI and machine learning techniques in hybrid movie recommendation system applying text-to-number conversion and cosine similarity approaches". *Journal of Computer Science and Technology Studies*, vol. 6, no. 1, pp. 94-102, 2024.

[7] M. Mouhiha, O. A. Oualhaj and A. Mabrouk. "*Combining Collaborative Filtering and Content Based Filtering for Recommendation Systems*". IEEE, United States, pp. 1-6, 2024.

[8] M. Loukili, F. Messaoudi and M. El Ghazi. "Machine learning based recommender system for e-commerce". *IAES International Journal of Artificial Intelligence*, vol. 12, no. 4, pp. 1803-1811, 2023.

[9] N. Cherkaoui, K. El Handri, M. D. Y. Tanoga, Y. El Hassani and A. Errafyg. "Consumer behaviour: Analysing marketing campaigns through recommender systems and statistical techniques". *Marketing and Management of Innovations*, vol. 15, no. 3, pp. 1-12, 2024.

[10] S. Ozturk, A. B. Ercan, R. Tugay and Ş. G. Öğüdücü. Enhancing Cross-Market Recommendation System with Graph Isomorphism Networks: A Novel Approach to Personalized User Experience. In: "*Proceedings of the 2024 9th International Conference on Machine Learning Technologies*", 2024.

[11] D. Patel, F. Patel and U. Chauhan. "Recommendation systems: Types, applications, and challenges". *International Journal of Computing and Digital Systems*, vol. 13, pp. 2210-2142, 2023.

[12] J. B. Schafer, D. Frankowski, J. Herlocker and S. Sen. Collaborative filtering recommender systems. In: P. Brusilovsky, A. Kobsa and W. Nejdl, editors. "*The Adaptive Web: Methods and Strategies of Web Personalization*". Springer, Berlin, Heidelberg, pp. 291-324, 2007.

[13] T. M. Mitchell. "*Machine Learning*". Vol. 1. McGraw-Hill, New York, 1997. Available from: https://www.pachecoj.com/courses/csc380_fall21/lectures/mlintro.pdf [Last accessed on 2025 Apr 19].

[14] F. Weiner, P. L. Teh, and C. B. Cheng. "*Systematic Review of Machine Learning in Recommendation Systems Over the Last Decade*". Springer Nature Switzerland, Cham, 2024.

[15] M. Chiny, M. Chihab, O. Bencharef and C. Younes. "*Netflix Recommendation System Based on TF-IDF and Cosine Similarity Algorithms*". no. BML, pp. 15-20, 2022. Available from: https://www.scitepress.org/Link.aspx?doi=10.5220/0010727500003101 [Last accessed on 2025 Apr 19].

[16] B. Sarwar, G. Karypis, J. A. Konstan and J. T. Riedl. Application of dimensionality reduction in recommender system-a case study. In: "*ACM WebKDD 2000 Web Mining for E-Commerce Workshop*". Boston, MA, USA, 2000.

[17] O. Kramer. K-Nearest Neighbors. In: "*Dimensionality Reduction with Unsupervised Nearest Neighbors*". Springer, Berlin, Heidelberg, 2013, pp. 13-23.

ORIGINAL RESEARCH ARTICLE UHD JOURNAL OF SCIENCE AND TECHNOLOGY

Utilizing Machine Learning Techniques for Cancer Prediction and Classification based on Gene Expression Data

Mariwan Mahmood Hama Aziz, Sozan Abdullah Mahmood

Department of Computer, College of Science, University of Sulaimani, Sulaymaniyah 46001, Kurdistan, Iraq

ABSTRACT

Cancer classification through genetic evaluation has become a hot topic among researchers. It holds the promise of delivering systematic, precise, and scientifically backed diagnoses for different types of cancer. Lately, several studies have delved into cancer classification by leveraging data mining techniques, machine learning algorithms, and statistical methods to thoroughly analyze high-dimensional datasets. Detecting cancer early by examining gene expression data is vital for providing effective patient care. Each sample in the Gene dataset usually includes a range of features, each representing a specific gene. In this paper, we propose a unique approach that utilizes DistilBERT, a distilled version of the Bidirectional Encoder Representations from Transformers, for cancer classification and prediction. In addition, our model integrates a self-attention mechanism in the transformer layers to enhance the model's focus on key features and employs an embedding layer for dimensionality reduction, improving the processing of gene statistics, preventing overfitting, and boosting generalization. We utilized datasets from important resources: The gene expression omnibus, which provided microarray records of lung and ovarian cancers, and the cancer genome atlas (TCGA), which offered RNA-Seq facts encompassing multiple most cancer types (breast invasive carcinoma, kidney renal clear cell carcinoma, colon adenocarcinoma, lung adenocarcinoma, and prostate adenocarcinoma). Our approach established excessive accuracy across all datasets, showcasing big upgrades in overall model performance compared to present strategies within the subject. The results underscore the ability to leverage transformer-primarily based architectures for strong cancer-type prediction and classification. Our approach achieved and improved exceptional accuracy compared to previous studies, with DS1: 97.56% for lung cancer, DS2: 100% for ovarian cancer, and DS3: 99.504% for the TCGA dataset.

Index Terms: Cancer Classification, Gene Expression Data, RNA-Seq, DNA Microarray, Bidirectional Encoder Representations from Transformers Model, Machine Learning, Pan-cancer, The Cancer Genome Atlas, DistilBERT

1. INTRODUCTION

Deoxyribonucleic acid, or DNA, stores genetic information needed by all living things to create, function, and develop. DNA is generally regarded as the blueprint of all living organisms since its components encode all of the information required to sustain life. Cancer is a complicated disease that stems from genetic mutations and unusual patterns of gene expression. These molecular shifts can throw off the normal functioning of cells, resulting in unchecked cell growth and the formation of tumors. Thanks to recent breakthroughs in gene expression profiling technologies, researchers can now assess the activity of thousands of genes all at once, offering crucial insights into how we diagnose, classify, and predict cancer outcomes [2], [3]. Cancer has become one of the most fatal illnesses globally, with an anticipated 9.7 million deaths from 20 million new cancer diagnoses in

Access this article online

DOI:10.21928/uhdjst.v9n1y2025.pp135-148

E-ISSN: 2521-4217
P-ISSN: 2521-4209

Corresponding author's e-mail: Mariwan Mahmood Hama Aziz, Department of Computer, College of Science, University of Sulaimani, Sulaymaniyah 46001, Kurdistan, Iraq. E-mail: mariwan.hamaaziz@univsul.edu.iq

Received: 05-04-2025 Accepted: 26-04-2025 Published: 02-06-2025

2022, according to the World Health Organization. Cancer is caused by the unrestrained proliferation of some abnormal cells, which divide and spread to other cells, multiplying malignant child cells. Men's most frequent cancers include lung, prostate, colorectal, and stomach [4]. Over the past two decades, health informatics research has focused on a variety of topics, including bioinformatics, cheminformatics, cancer prediction, and others [5].

Gene expression is the method by which the knowledge stored in DNA is transformed into instructions for producing proteins or other substances. It starts with the transcription of DNA into messenger RNA, which is then translated into proteins. Gene expression analysis is used to analyze the order of genetic modifications occurring under specific conditions in tissue or a single cell [6], [7]. A new technique for studying the expression of several genes at once is microarray technology. It entails positioning thousands of sequences of genes on a glass slide known as a "gene chip" in specific locations. The gene chip comes into contact with a sample of DNA or RNA. Measured light is produced by complementary base pairing between the sample and the gene sequences on the chip. Genes expressed in the sample are identified by regions of the chip that emit light. Each row in a tabular representation of a microarray gene expression data set corresponds to a single gene, each column to a sample or time point, and each matrix entry represents the measured expression level of a specific gene in a sample [8]–[10]. By offering more normalized and less noisy data for classification and prediction, RNA-Seq is a novel and well-liked method for finding new transcripts and isoforms. Finding the genes that are differentially expressed in a body or identifying changes in genes at various levels is the primary purpose of transcriptome profiling. RNA sequencing allows for both identification and quantification in one location. RNA-Seq data are widely available from various databases that can be used for cancer prediction and classification [11].

Machine learning (ML) techniques have recently been utilized to analyze microarray datasets for the categorization of cancer. One useful method for diagnosing cancer is to use the gene expressions found in microarray datasets. Several feature selection techniques have been used to identify the most important properties of malignant microarray datasets to enhance the performance of these widely used ML algorithms [6]. Notably, several innovative algorithms have surfaced that have demonstrated encouraging outcomes across a range of fields [5]. A subfield of artificial intelligence called ML gives computers the ability to learn from training data, identify patterns in data, and make predictions on their own that get better over time without explicit programming. Numerous classification techniques were developed in the ML field, and many of them were applied to the categorization of cancer [12].

In this paper, we advocate a technique called DistilBERT, which is a distilled version of the Bidirectional Encoder Representations from Transformers (BERT) model that retains 97% of BERT's language understanding power while being lighter, faster, and smaller. DistilBERT was first presented by Hugging Face and is designed especially for tasks requiring less processing power [13]. The BERT version, added by Wu *et al.* (2024), is a groundbreaking deep learning model designed for herbal language processing that has 110 million parameters for the base version and 340 million parameters for the large version. Unlike conventional models that study textual content input in a unidirectional manner. It uses a transformer structure that reads the input text bi-directionally. This lets it recognize the context of a phrase based totally on both its left and proper environment, imparting deeper semantic knowledge [14]. BERT's transformer-primarily based architecture has been adapted for diverse fields past textual content processing, such as bioinformatics and computational biology. In those packages [15]–[17]. DistilBERT, such as the BERT model, uses the same structure but is compressed to reduce model size, holding most of BERT's overall performance with fewer parameters – approximately 60% of the size of BERT (66 parameters), making it quicker and more efficient, and providing quicker predictions with high performance [18].

Two forms of gene expression datasets from diverse sources are used in this study to evaluate the efficacy of the recommended method, and the selected data do not achieve the high results with the previous model. The gene expression omnibus (GEO) provides microarray datasets, which include samples of ovarian and lung cancer. The availability and dependability of these microarray datasets, which provide a photo of gene activity, have made them famous for being used in most cancer studies [11], [19]. The 2nd set of statistics is derived from the cancer genome atlas (TCGA), a comprehensive RNA-Seq dataset that consists of facts on numerous cancer types, including prostate adenocarcinoma (PRAD), lung adenocarcinoma (LUAD), colon adenocarcinoma (COAD), kidney renal clear cell carcinoma (KIRC), and breast invasive carcinoma (BRCA). We can very well evaluate the adaptability and efficacy of the DistilBERT model across various gene expression technologies by way of the utilization of each microarray and RNA-Seq information [20], [21].

The structure of this paper is prepared as follows: The Methods section offers a complete evaluation of the datasets and pre-processing strategies employed. The architecture and implementation phase into the version of DistilBERT for numerical input and its integration into the cancer category framework. Finally, we gift the experimental results, comparing our method with existing today's models, observed with the aid of a dialogue at the implications and capacity applications of these studies in customized oncology and medical decision aid.

2. LITERATURE REVIEW AND PROBLEM STATEMENT

This section reviews key research in cancer classification, highlighting the transition from traditional methods to innovative approaches of deep learning and optimization techniques. It showcases studies that utilize gene expression data and delves into how metaheuristic algorithms have been employed to enhance feature selection and boost model performance.

2.1. ML-Based Methods

Tabassum *et al.* (2024) [3]. Proposed an ensemble learning approach that uses a bagging-based multilayer perceptron's and mutual information for feature selection to classify cancer from high-dimensional gene expression data. The method was applied to different cancer types, demonstrating its effectiveness in handling high-dimensional data and achieving varying levels of accuracy across several datasets. In this study, AbdElNabi *et al.* (2020) [4]. Introduced an intelligent decision support system for cancer classification using gene expression data from breast and colon cancers. Their method combines information gain (IG) for initial feature selection, Grey Wolf Optimization for further dimensionality reduction, and a support vector machine (SVM) for classification. Applied to microarray datasets, the approach effectively handled high-dimensional data and achieved strong classification performance, demonstrating its stability and reliability in early cancer diagnosis. Other studies by Guyon *et al.* (2002) [22]. Integrating recursive feature elimination (RFE) with SVM. The RFE method, used for gene choice, finished with incredible accuracy, and SVM for cancer classification consisted of 98% on leukemia datasets. A recent study introduced a two-phase hybrid feature selection method by Ali and Saeed (2023) [6]. Combining filter techniques (IG, gain ratio, Chi-squared) with genetic algorithms (GA) to improve cancer classification. The approach was tested using SVM, Naive Bayes, k-nearest associates (KNN), Decision Tree, and random forest (RF) on microarray datasets for breast, lung, Central Nervous System, and brain cancers. The GA step further refined features selected by filters, enhancing overall classification performance, Wei *et al.* (2023) [23]. Emphasized the importance of feature extraction and selection in high-dimensional gene expression data. They applied methods such as methods like principal component analysis (PCA), IG, and GA were broadly followed. A look at applying PCA with numerous classifiers, which includes choice trees (DT), SVM, and RF, performed variable outcomes, emphasizing the importance of effective function extraction in optimizing model overall performance, Li *et al.* (2020) [24]. Carried out an extensive study on pan-cancer classification, utilizing TCGA RNA-seq gene expression data from 31 different tumor types. They employed ML techniques to pinpoint groups of distinguishing genes that could differentiate between these tumor types with an impressive accuracy of over 90%. The research also delved into sex-specific variations in gene expression, underscoring the promise of certain biomarkers for tumor diagnosis and tailored treatment approaches. In this approach, García-Díaz *et al.* (2022) [25]. Proposed unsupervised studying strategies have additionally been explored for multiclass cancer classification. A look at employing an extreme learning machine with a genetic grouping algorithm completed a median accuracy of 98.8% for breast, kidney, and prostate cancers, demonstrating the feasibility of unsupervised techniques for high-dimensional data. In addition, Chen (2022) [26]. Presented ML models, which include SVM, linear discriminant analysis (LDA), and KNN, have also been explored for multi-cancer datasets, consisting of brain, prostate, and colon cancers. These fashions did F-scores above 80% and furnished insights into feature screening techniques for dealing with high-dimensional gene expression data. In another study, gene choice strategies have additionally been tailored for cancer classification by AlShamlan and AlMazrua (2024) [5]. An examination leveraging Harris Hawks Optimization and KNN completed perfect typing for colon tumors and leukemia datasets. These effects spotlight the promise of biostimulator algorithms in identifying biologically applicable gene markers. Mukhopadhyay *et al.* (2023) [12]. Proposed discriminant analysis (LDA) combined with RF is explored for excessive-dimensional microarray gene expression facts. The study finished with accuracies of 96% for breast cancer, 98% for most colon cancers, and 99% for most prostate cancers, demonstrating the effectiveness of dimensionality discount strategies in improving category overall performance for multi-cancer datasets. Brought a bendy category framework for cancer gene expression profiles by Hijazi and Chan (2013) [20]. Utilizing ML models, such as DT, RF, and KNN, they have a look at implementing more than one characteristic

choice strategy (filter out, wrapper, and embedded) to datasets together with leukemia, colon, and prostate cancer, showcasing the adaptability of ML frameworks throughout extraordinary cancer kinds.

2.2. Deep Learning and Hybrid Approaches

Similarly, another study by Das *et al.* (2023) [27]. Use CNN, LSTM, and hybrid architectures, such as DCNN-GRU with enhanced chimp optimization algorithms to classify cancer using microarray data. The researchers tested these models on datasets that included various subtypes such as brain, breast, prostate, colon, and leukemia. These approaches leverage deep learning capacity to capture complicated styles, supplying sturdy consequences in gene expression-based cancer detection, Yaqoob *et al.* (2023) [28]. Proposed Recent research has furthermore delivered hybrid algorithms to beautify most cancer classes; integrated ML classifiers for breast cancer classification, such as KNN, SVM, and Naive Bayes, with the sine cosine and cuckoo search algorithm (SCACSA) brought about high performance in breast cancer types, outperforming traditional techniques. The study presents limitations that are important to consider. The SCACSA method relies on the quality and size of the dataset used for validation. On the other hand, Tarek *et al.* (2016) [29]. The KNN set of rules has also proven promise in most cancer predictions. Have a look at applied wrapper, filter out, and embedded feature choice methods with microarray datasets for leukemia, colon, and breast cancers, achieving accuracy rates of 99% and 100%, respectively, showcasing the adaptability of KNN across exclusive cancer datasets. Rukhsar *et al.* (2022) [2]. Introduced a deep-learning framework for classifying multiple types of cancer using RNA-Seq gene expression data. They took the complex, high-dimensional gene data and converted it into 2D images through processes, such as normalization and zero-padding. Then, they employed eight different deep learning algorithms, including CNN, to extract features and categorize samples from five distinct cancer types. Their experiments, which involved various data splits and k-fold cross-validation, showed that CNN outshone the other models in terms of classification performance, achieving a high accuracy of 97%, Mohammed *et al.* (2023) [11]. Implemented hybrid stacking ensembles, which have furthermore proven powerful. For instance, employing 1D-CNN and LASSO with TCGA datasets yielded accuracies of 99.54 % for full datasets and 98.62% for reduced datasets, demonstrating the performance of deep learning with dimensionality reduction strategies. Some studies by Sucharita *et al.* (2024) [19]. Have centered on enhancing cancer type classification through deep learning improvements. For example, a hybrid version combining exponential sigmoid-deep notion networks and ranking methods carried out accuracies of 85–95% throughout seven cancer kinds, including leukemia and ovarian cancers, illustrating the potential of deep belief networks in gene expression evaluation. Aburass *et al.* (2024) [30]. Introduced a hybrid ML model combining CNN, LSTM, and GRU architectures for gene mutation category execution, achieving 80.6% accuracy, and suggesting opportunities for additional optimization in hybrid frameworks. Despite these improvements, challenges persist in attaining regular generalization throughout datasets and addressing the computational complexity of high-dimensional records evaluation. In previous work by Thakur *et al.* (2024) [21]. Multi-cancer analysis has, moreover, benefited from advancements in ML. A comprehensive genomic pan-cancer category using TCGA datasets was completed with 90% accuracy through integrating GA, demonstrating the value of function choice in large-scale genomic statistics evaluation. Another effort mixed RNN-CNN architectures with bottleneck function extraction, attaining accuracies of 97.8% for breast cancer and 99.4% for prostate cancer. In this study, Surbhi Gupta *et al.* (2023) [10]. Posited deep studying strategies continue to be pivotal for various cancer types. Deep learning on RNA sequence datasets was examined for breast, lung, kidney, prostate, and colon cancers. Although unique accuracy values are no longer certain, these studies demonstrate the strong potential of deep learning architectures in managing complex datasets, reinforcing their relevance in modern-day cancer studies. An innovative graph convolutional network (GCN) was applied to TGCA datasets by Martínez Logreira (2020) [31]. Attaining approximately 52% accuracy for pan-cancer evaluation. Although the performance became modest, this observation highlighted the potential of graph-based procedures for shooting complex relationships in genomic records. Table 1 provides a precise view of the literature discussed above.

2.3. Limitations of Existing Work

Although there are significant advances in cancer classification using gene expression data, several recurring challenges continue to limit the effectiveness and scalability of existing methods. Key limitations identified in recent studies include.

- Lack of generalization: A lot of models are trained and fine-tuned on specific datasets, but they often skip validation on external or diverse datasets.
- Dataset dependency: When sample sizes are small or when there's a heavy reliance on just microarray or RNA-Seq data, it limits how well these models can apply to a wider range of cancer types.
- Computational cost: Methods that rely on optimization, such as GA and deep learning frameworks, tend to be

TABLE 1: Comparative review of literature

References	Model	Feature extraction	Dataset	Year
[3]	Multilayer perceptron's (MLPs)	Mutual information algorithm	Microarray	2024
[29]	k-nearest neighbors (KNN) algorithm	Wrappers, Filters, Embedded methods	Microarray	2016
[27]	CNN, LSTM, DCNN, GRU, PSCS-DL, CSSMO-DL	ECO algorithm	Microarray	2024
[4]	SVM	Information gain (IG)	Microarray	2020
[6]	SVM, NB, KNN, DT, RF	IG, information gain ratio, and Chi-squared	Microarray	2023
[2]	CNN	Deep learning (DL)	RNA-Seq data	2022
[25]	Extreme learning machine (ELM)	Grouping genetic algorithm (GGA)	RNA-Seq data	2020
[11]	1D-CNN	LASSO	(TCGA)	2022
[26]	SVM, LDA, or KNN	Feature screening	Gene expressions	2022
[22]	SVM with RFE	Recursive feature elimination (RFE)	Leukemia data	2002
[19]	Exponential sigmoid-deep belief network (ES-DBN)	Feature ranking (CM-CRO,)	Microarray Data	2024
[5]	(KNN), (SVM),	Harris hawks optimization (HHO)	Microarray Data	2024
[12]	Linear discriminant analysis (LDA) and (RF)	Linear discriminant analysis (LDA)	Microarray	2024
[30]	LSTM, LSTM, CNN, GRU	Not mention	Cancer Treatment dataset	2024
[28]	SVM, KNN, NB	(SCACSA)	Microarray Data	2024
[23]	DT, SVM, RF, NB, Neural network, KNN	Principal component analysis (PCA)	Microarray Data	2023
[20]	DT, SVM, RF, KNN, bagging,	Filter, wrapper, and embedded methods	Microarray Data	2013
[24]	KNN	(GA)	(TCGA)	2017
[21]	RNN-CNN	Sandwich stacked method based on VGG16 and VGG19 pre-trained models	Gene expression data	2023
[10]	Deep learning	Not mentioned	RNA sequence dataset	2022
[31]	Graph convolutional network (GCN)	Genetic Algorithms (GA)	The TGCA dataset	2020

ECO: Enhanced chimp optimization

resource-heavy, making them less ideal for real-time applications or environments with limited resources.

- Manual or static feature selection: Many studies stick to traditional feature selection techniques that need manual adjustments and don't adapt on the fly during training.
- Limited data pre-processing and hyperparameter tuning: Some methods fall short on having effective pre-processing steps or optimized hyperparameter choices, which can hurt their overall performance.
- Limited use of advanced models: There's a noticeable lack of exploration into transformer-based or graph-based neural networks in many studies, even though these could do a better job of capturing complex relationships between genes.

2.4. Problem Statement

Cancer diagnosis remains a critical challenge in healthcare, where early and accurate detection is essential to improving outcomes and reducing mortality. Traditional methods often fall short in handling the complexity of gene expression data, and many ML approaches struggle with generalizability, static feature selection, and dataset-specific tuning. To address these issues, this study introduces a DistilBERT-based model with a self-attention mechanism that dynamically identifies significant gene features during training. This approach enhances accuracy, reduces manual pre-processing, and offers a scalable solution for classifying multiple cancer types using high-dimensional gene expression data.

3. MATERIALS AND METHODS

The main steps in developing this research for cancer classification using gene expression include data collection, data pre-processing, gene selection using the self-attention mechanism, and finally classification using the DistilBERT model. Fig. 1 describes the processing steps of the proposed methodology; each step is briefly described next.

3.1. Data Collection

To determine the effectiveness of our DistilBERT model for cancer classification across different cancer types, we utilized publicly available gene expression datasets from two open sources platform, such as the GEO and TCGA. These data are not re-identifiable and have been released under a license that prohibits their use for commercial purposes only. They were used in a way that matches the requirements, that is, under a subscription based upon the terms and conditions established by both GEO and TCGA. The study does not need the formal approval of the Institutional Review

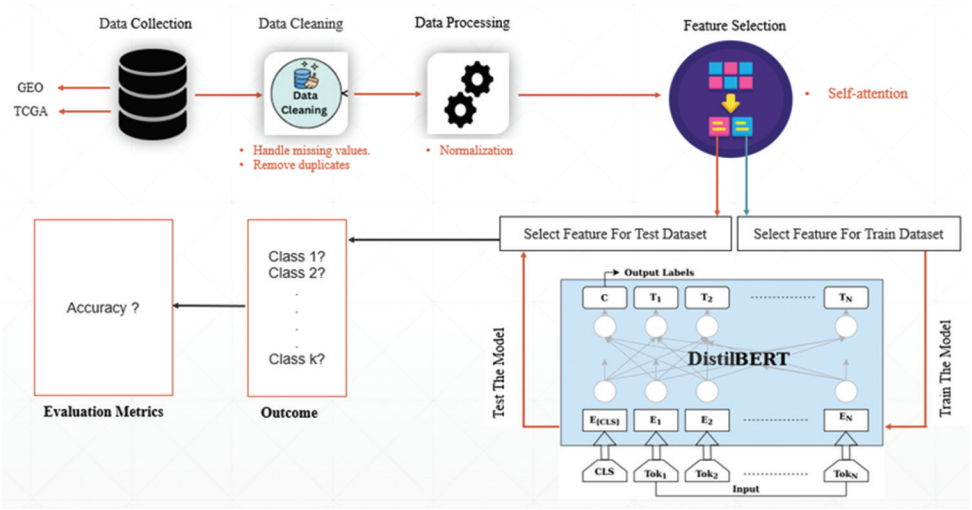


Fig. 1. Steps of the proposed methodology for cancer classification.

Board as it is public data and it does not contain any private information that has an identifiable person. We specifically selected datasets that had presented challenges to previous models, aiming to demonstrate the potential of our approach and achieve better results.

The GEO is managed by the National Center for Biotechnology Information. It is a repository for high-throughput gene expression and other functional genomics data. From GEO, we downloaded two high-dimensional microarray datasets: one for lung cancer and another for ovarian cancer [6] (as detailed in Table 2). The lung cancer dataset contains 203 instances with five classes and 12,600 features, or genes. The samples in the lung cancer dataset are classified as belonging to four classes of lung tumors: Small cell lung cancer (6 samples), adenocarcinoma (139 samples), normal lung (17 samples), squamous cell carcinoma (21 samples), and pulmonary carcinoid (20 samples) [2]. The ovarian cancer dataset includes 253 samples with two classes and 15,154 genes. The ovarian cancer dataset is labeled with a normal class (91 samples) and with cancer classes (162 samples) [3].

The second data source were TCGA, More than 20,000 primary cancers and matched normal samples from 33 different cancer types were molecularly characterized by the groundbreaking Cancer Genome Atlas (TCGA) program. Beginning in 2006, this collaborative effort between NCI and the National Human Genome Research Institute brought together scientists from various institutions and disciplines. In this source, we downloaded the RNA-Seq gene expression data from Pan-Cancer Atlas (https://portal.gdc.cancer.gov/) using the R statistical application version 3.6.3 by the TCGAbiolinks package [2], [11]. The dataset contains 801 instances or samples and 20,531 features or genes from the top five common cancer types, including BRCA, KIRC, COAD, LUAD, and PRAD [2], [11], [24]. Each sample has 20,532 gene sequences. The dataset's cancer classes are denoted by the following codes: 0, 1, 2, 3, and 4 for PRAD, LUAD, BRCA, KIRC, and COAD. Out of a total of 801 samples, the BRCA class has 300 samples, clear cell carcinoma (KIRC) has 146, LUAD has 141, COAD has 78, and PRAD class has 136 samples [21], [24]. As shown in Table 3, after downloading, we combine each type of cancer to make a unified, large-scale dataset for training and evaluating our model, aiming for a more generalized and accurate approach to cancer classification and prediction across multiple cancer types.

TABLE 2: Description of the high-dimensional microarray datasets used in this study

Dataset	No. of features	No. of instances	No. of classes
DS1: Lung cancer	12,600	203	5
DS2: Ovarian cancer	15,154	253	2

TABLE 3: Description of the DS3: Pan-cancer datasets used in this study

Dataset	No. of features	No. of instances
BRCA	20,532	300
KIRC	20,532	146
LUAD	20,532	141
COAD	20,532	78
PRAD	20,532	136

3.2. Data Processing

Before using a ML model, the selected datasets must be properly processed and processing raw gene expression data can be challenging due to its varied range. Several common procedures are taken during the pre-processing stage, including Data Cleaning. Normalization and feature selection.

3.2.1. Data cleaning

To ensure the quality and reliability of our datasets for effective model training, a data cleaning phase was performed. This involves identifying and dealing with different facets of data quality, including missing values, duplicates, inconsistencies, and outliers, which can lead to poor performance and interpretability of ML models. Removing missing values and duplicates is a very important step toward statistics cleaning to ensure the quality and abundance of the dataset, substituting missing values with statistical measures such as mean, median, or mode such as mean, median, or mode as shown in Tables 4 and 5. Similarly, duplicate records in a dataset can distort evaluation and version performance. Identifying and getting rid of duplicates guarantees statistics integrity and decreases redundancy [3], [27], [32].

3.2.2. Normalization

To ensure that all gene expression features contributed equally to the model training process, we applied normalization using the StandardScaler technique. The goal of normalization is to convert the values of numeric columns in the dataset to a common scale, which improves both the performance and accuracy of your model without distorting value ranges or losing any information [33]. We specifically employed StandardScaler, which centers the data around a mean of 0 and a standard deviation of 1. Figure 3 shows the data before normalization, and Figure 4 demonstrates the data after normalization. StandardScaler enhances version education balance by preventing features with larger scales from dominating other [34]. The scikit-learn (sklearn) library in Python includes the StandardScaler implementation. Fig. 2 is the form of the script we used when StandardScaler was implemented before data splitting was done.

```
scaler = StandardScaler ()
X_scaled = scaler.fit_transform(X)
```

Fig. 2. Standardscaler Implementation [1].

We use StandardScaler in normalization, and the equations (1), (2), and (3) represent the metaethical formula of standardization, mean, and standard deviation. Where X is the original value of the feature, N is the total number of values in the dataset, μ is the mean of the feature, and σ is the standard deviation of the features [35]–[37].

$$\text{X standardization} = \frac{x - \mu}{\sigma} \tag{1}$$

$$\text{Mean } \mu = \frac{1}{N}\sum_{i=1}^{N} Xi \tag{2}$$

$$\text{Standard Deviation } \sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(Xi - \mu)2} \tag{3}$$

3.2.3. Feature selection

To identify the most relevant gene expression features for reliable cancer classification, we employed the inherent self-attention mechanism within the DistilBERT model for feature selection. This training process, learning their significance without explicit pre-processing, evaluates the relationships and dependencies among capabilities, assigning attention weights that reflect their importance in the context of the given project. Unlike traditional feature selection strategies, which necessitate either manual guidance or algorithmic

	AFFX-MurIL2_at	AFFX-MurIL10_at	AFFX-MurIL4_at	AFFX-MurFAS_at	AFFX-BioB-5_at	...	105_at
0	-18.600	10.54	0.010	19.440	-16.980	...	1.630
1	9.120	9.12	10.180	29.290	-4.680	...	10.180
2	-2.175	-2.21	-0.060	6.320	-1.775	...	1.745
3	-1.540	21.75	5.835	23.815	-24.785	...	10.355
4	-9.070	3.08	-1.980	17.260	-10.090	...	-10.090
..	...	...	...	...	...	...	...
198	35.140	106.16	52.280	65.340	27.790	...	48.200
199	-21.150	-31.20	-11.820	8.280	-24.740	...	-3.210
200	26.900	10.44	18.230	33.830	-11.220	...	6.970
201	23.800	29.14	31.800	65.610	4.240	...	26.470
202	-18.370	-1.03	-8.260	27.150	-23.430	...	-4.640

[203 rows x 12600 columns]

Fig. 3. Example of data before normalization.

pre-processing to pinpoint and eliminate irrelevant features, it dynamically learns which capabilities (genes) are maximally relevant for distinguishing between different cancer types. [38]–[42]. This concept is mathematically captured by what's known as scaled dot-product attention.

$$\text{Attention Score}(Q, K, V) = \text{softmax}\left(\frac{Qk^T}{\sqrt{d_k}}\right) V \quad (4)$$

Where Q, K, and V are the query, key, and value matrices derived from gene expression embeddings. By focusing on these dynamically identified key features, the model can potentially achieve better performance and generalization [39].

3.2.4. Data splitting

To train our DistilBERT model and see how well it performs on new, unseen data, we split each of our datasets into training and testing sets. Training data contained up to 80% of the overall dataset, allowing it to learn the patterns in the gene expression profiles associated with different cancer types, whereas test data represented up to 20%. As a result, the DistilBERT models were used to classify the cancer types.

```
[[-0.52270655  0.24721653 -0.02648141 ... -0.40856243 -0.9889013
   0.18736746]
 [ 0.82877923  0.18378861  0.50429813 ... -1.27053002  0.60904913
   1.18361247]
 [ 0.27809265 -0.3222947  -0.03013476 ... -0.48253143  0.41890506
   0.07492172]
 ...
 [ 1.69564132  0.24274977  0.92443336 ... -1.35286656  1.40594247
   0.01376094]
 [ 1.54450113  1.07803291  1.63266132 ...  1.90043074  0.55746503
   2.20064867]
 [-0.51149292 -0.26958699 -0.4580986  ...  0.93894528  0.82972729
  -1.08817267]]
```

Fig. 4. Example of data after normalization by Standardscaler.

4. PROPOSED CLASSIFICATION MODEL

Our proposed classification model uses a modified DistilBERT architecture to classify cancer types based on high-dimensional numerical gene expression datasets. This model is intended to efficiently process and evaluate input information to accurately forecast the cancer class and use our model with different types of cancer and achieve the highest accuracy.

4.1. DistilBERT

DistilBERT is a streamlined version of the BERT model, as shown in Fig. 5. It starts by taking in high-dimensional numerical inputs that represent gene expression levels. The first step is an embedding layer that uses a linear transformation (nn.Linear) to shrink the input dimensions down to 768, setting the stage for the next steps. To enhance training stability and improve generalization, a Group Normalization (Group Norm) layer is applied right after the embedding layer, ensuring that the input to the Transformer is appropriately normalized. After that, the model goes through a Transformer block made up of six stacked layers. Each of these layers features a self-attention mechanism, allowing the model to hone in on the most significant gene features by assigning varying weights across the input sequence. A special [CLS] token is added at the beginning, and it gets refined through these layers to serve

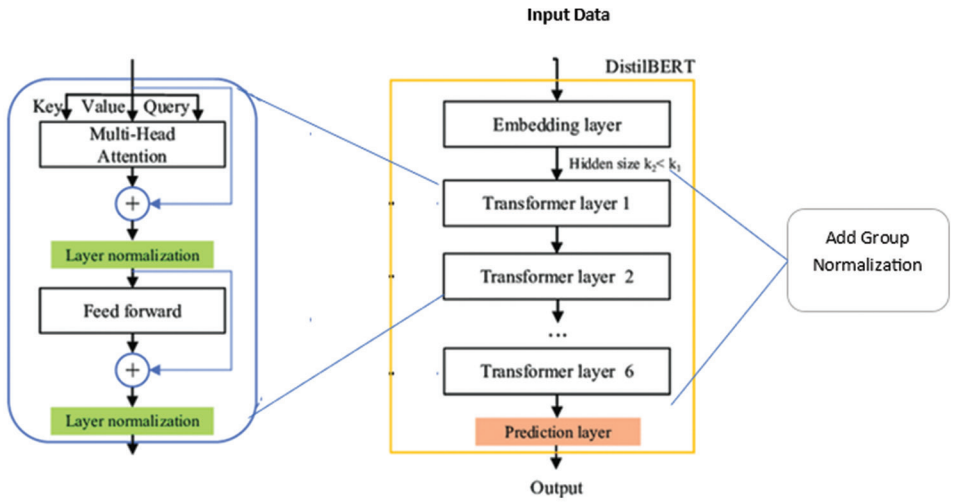


Fig. 5. Workflow of the DistilBERT model approach used as an ensemble classifier.

as a global summary of the entire sample. Following the self-attention process, a feed-forward network (FFN) – which is a fully connected network – adds non-linearity, enriching the feature representations and boosting the model's learning capabilities. Before reaching the final classification stage, a dropout layer with a rate of 0.2 is applied to the [CLS] token to help prevent overfitting and improve the model's generalization. Finally, a fully connected classification layer translates the 768-dimensional [CLS] representation into the output space, and the results are processed through a SoftMax function to generate the final class probabilities. One of the key strengths of the proposed study lies in its capability for generalization. By using a flexible transformer-based architecture, the version may be tailored to categorize extraordinary sorts of cancer with minimal changes. This generalizability becomes evident within the steady overall performance across various datasets examined throughout the study.

4.2. Training

During the training step, the model uses optimizers, such as AdamW and SGD with a learning rate of 2×10^{-5} and a weight decay of 1×10^{-3} and uses cross-entropy loss for classification, which allows the model to learn the relationships between the input features and the cancer classes.

4.3. Experimental Setup

The model we proposed was built using Python 3.11.5 and Visual Studio Code. We made use of several essential libraries. We preprocessed and normalized different types of cancer from two open sources. For cancer classification, we fine-tuned DistilBERT, a streamlined transformer model. This model features 6 transformer layers, 12 attention heads, and 768 hidden units and can handle a maximum input length of 512. All our experiments were conducted on a Windows 10 machine equipped with an Intel Core i7-6820HQ CPU, 16 GB of RAM, and an NVIDIA GeForce RTX 4060 GPU.

5. RESULTS ANALYSIS

The general efficacy of the proposed DistilBERT-based complete variant for most cancer classes was examined using different datasets, including lung cancer, ovarian cancer, and TCGA datasets. The results show the model's ability to efficiently handle high-dimensional gene expression data and achieve exceptional class accuracy across several datasets. The results are evaluated using accuracy performance metrics. In our experiments with the DistilBERT model, we explored various batch sizes and used the AdamW and SGD optimizers on all datasets. For the lung cancer dataset, our approach achieved an exceptional accuracy improvement of 97.56% with the AdamW optimizer and a batch size of 16, outperforming other batch sizes and SGD, making it the best choice for this task. Conversely, the SGD optimizer showed optimal performance with batch sizes of 32 and 128, as shown in Table 6. Fig. 6 displays the highest accuracy and loss attained by the model for lung cancer datasets, and Fig. 7 shows the confusion matrix model.

For the ovarian cancer dataset that contains 15,154 genes with two classes, our model gets 100% accuracy with the AdamW optimizer across all batch sizes; with the SGD, the results are represented in Table 7 with different batch sizes. Fig. 8 displays the highest accuracy and loss attained by the model for ovarian cancer datasets, and Fig. 9 shows the confusion matrix model.

For the TCGA dataset that includes five types of cancers and contains 20,532 genes for each type, we use DistilBERT with

TABLE 4: Example of data with miss values and duplicate rows

Sample ID	Feature1	Feature2	Feature3	Feature4
Sample_0	0.0	2.017209	3.265527	5.478487
Sample_1	0.0	0.592732	1.588421	7.586157
Sample_2	Nan	NaN	2.3271	6.881787
Sample_3	0.0	3.511759	4.327199	6.881787
Sample_4	0.0	3.511759	4.327199	6.881787

TABLE 5: Example of data after handle missing values with remove duplicate rows

Sample ID	Feature1	Feature2	Feature3	Feature4
Sample_0	0.0	2.017209	3.265527	5.478487
Sample_1	0.0	0.592732	1.588421	7.586157
Sample_2	0.0(Mean)	2.408865(Mean)	2.3271	6.881787
Sample_3	0.0	3.511759	4.327199	6.881787

TABLE 6: Using different batch sizes with AdamW and SGD optimizers for the lung cancer dataset

Batch size	Optimizer	Accuracy %
16	AdamW	0.9756
32		0.926
64		0.9268
128		0.9268
16	SGD	0.9024
32		0.92
64		0.878
128		0.92

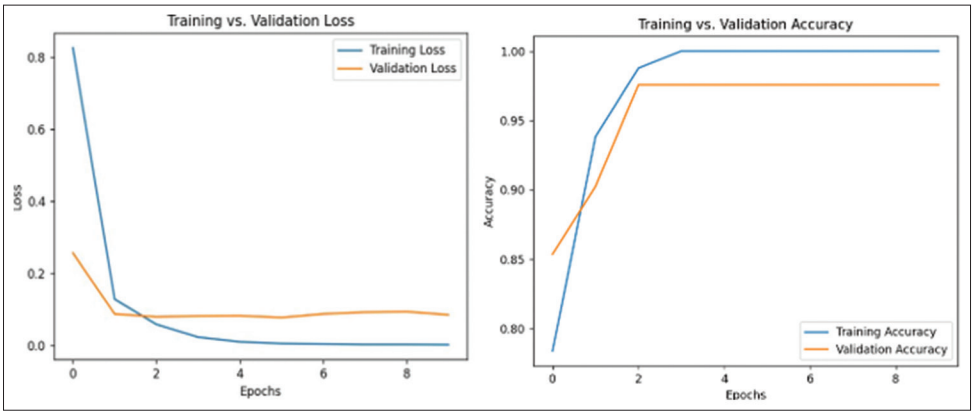


Fig. 6. Accuracy and loss for DistilBERT model for lung cancer dataset.

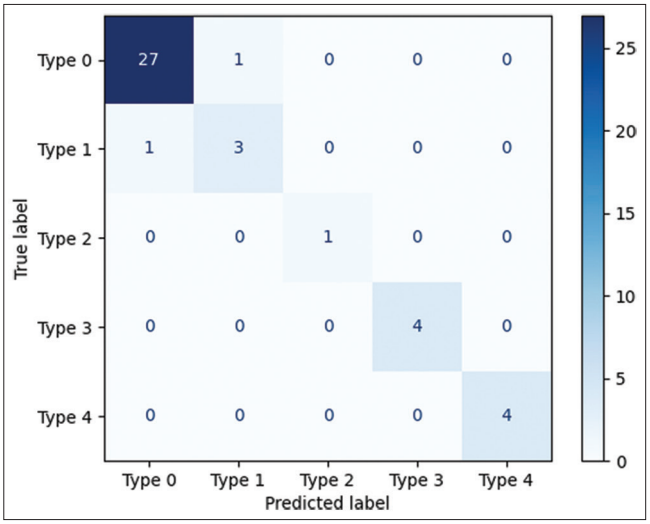


Fig. 7. Confusion matrix model for lung cancer dataset.

AdamW and SGD optimizers. Our model achieves 99.5% accuracy with the AdamW optimizer with a batch size of 16, as shown in Figs. 10 and 11 shows the confusion matrix. For the SGD optimizer, the result is presented in Table 8 with different batch sizes for both optimizers.

6. DISCUSSION

This paper proposed an approach by utilizing ML techniques for cancer prediction and classification based on gene expression data. The DistilBERT model was applied to classify gene expression datasets in bioinformatics because it can find complex non-linear relationships between the inputs and the outputs and is effective for large datasets. As discussed, it above indicates that DistilBERT model architectures could be used with two types of optimizers,

TABLE 7: Using different batch sizes with SGD for ovarian cancer

Batch size	Optimizer	Accuracy %
16	SGD	100
32		100
64		0.9804
128		0.9412

TABLE 8: Using different batch sizes with AdamW and SGD optimizer

Batch size	Optimizer	Accuracy %
16	AdamW	0.995
32		0.9937
64		0.9804
128		0.98
16	SGD	0.987
32		0.9813
64		0.9565
128		0.9255

SGD and AdamW, for cancer classification, as they show a confident result. The results indicate that model architectures performed well. The AdamW optimizer reached higher results across different batch sizes with model architecture, showing a better choice as its accuracy reached 97.56% with a batch size of 16, for the lung cancer dataset, while the accuracy dropped to (92.6%, 92%, and 92.68%) with batch sizes of 32, 64, and 128. For ovarian cancer the model performed well, and the accuracy achieved (100%) with the AdamW optimizer was the same with all different batch sizes. For the TCGA dataset, the accuracy reached (99.5%) with a batch size of 16, which is higher than the results for batch sizes of 32, 64, and 128 (99.37%, 98.04%, and 98%). On the other hand, the SGD optimizer produced the results in different batch size values, which are (16, 32, 64,

TABLE 9: Comparison of our proposed model with some existing works used in this field

Dataset	Existing work			Accuracy of our proposed model
	Method	Accuracy (%)	Year	
DS1: Lung Cancer	SVM, KNN, DT, RF [6]	94.09–97.04	2023	97.56%
	CNN [2]	97	2022	
	ES-DBN [19]	94.5454	2024	
	LDA and RF [12]	95	2024	
	RNN-CNN [21]	0.97	2023	
	SVM, RF, MLP, SMO [7]	93, 96, 86.6, 91	2022	
DS2: Ovarian Cancer	MLPs [3]	98	2024	100%
	1D-CNN [11]	98.62	2022	
	ES-DBN [19]	95.7746	2024	
	MI, GA, SVM [7]	80–98	2022	
DS3: TCGA Dataset	1D-CNN	98.62	2022	99.504%
	ELM [25]	98.81	2020	
	KNN [24]	90	2017	
	GCN [31]	52	2020	
	CNN [2]	97	2022	

SVM: Support vector machine, LDA: Linear discriminant analysis, MLPs: Multilayer perceptrons, KNN: k–Nearest Neighbors, ELM: Extreme learning machine, GCN: Graph convolutional network, ES-DBN: Exponential sigmoid-deep notion networks

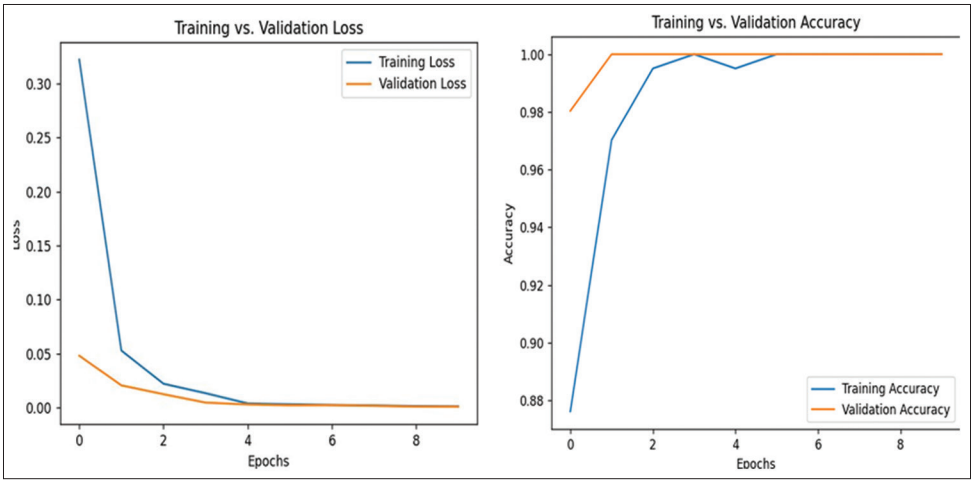


Fig. 8. Accuracy and loss for DistilBERT model for ovarian cancer dataset.

and 128). The accuracy was (90.24%, 92%, 87.8%, 92%) for lung cancer, but only in 16 and 32 did it reach the result of (100%), for ovarian cancer with model architecture. For the TCGA dataset, the accuracy is decreased with SGD for model architecture, as shown in Table 8. This suggests that batch sizes of 16 and 32 are the most effective for achieving optimal performance. This experiment discovered that the smaller batch size of 16 worked better because it allowed for improved generalization from the noisier updates, and the AdamW optimizer with batch 16 is more proper for cancer classification, as it got a high accuracy in almost all the tests compared to the SGD optimizer. The consistent 100% accuracy on the ovarian dataset with AdamW was unexpected, suggesting a highly separable gene expression

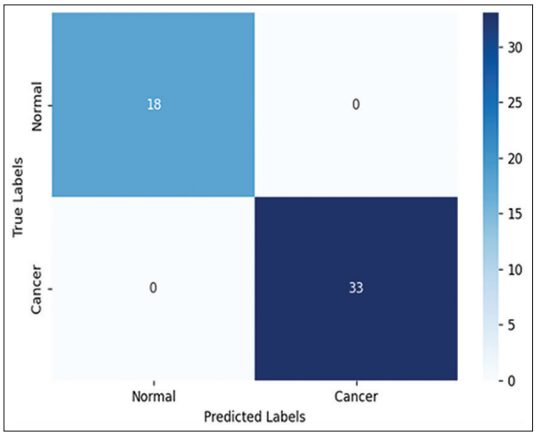


Fig. 9. Confusion matrix model for ovarian cancer dataset.

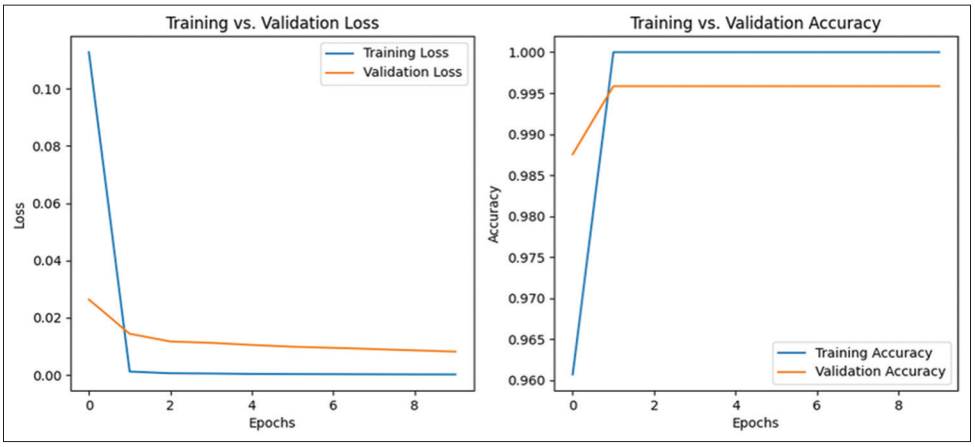


Fig. 10. Accuracy and loss for the DistilBERT model with the TCGA dataset.

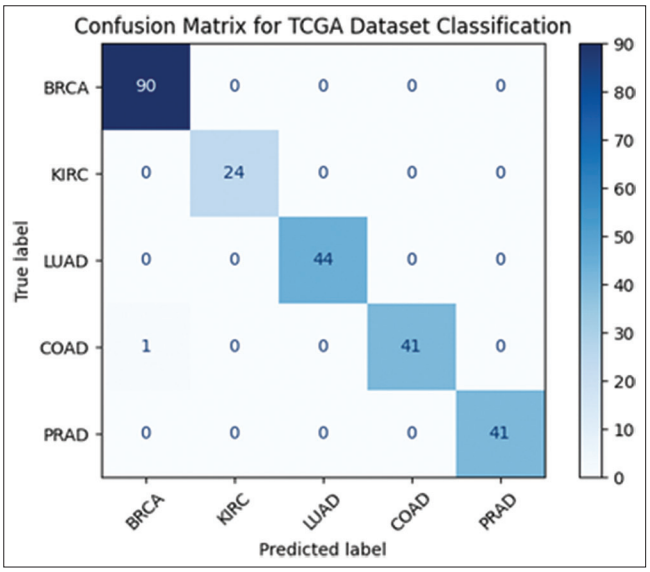


Fig. 11. Confusion matrix of the DistilBERT model with the TCGA dataset.

profile that simplifies classification. When we compared it to previous studies, the proposed model achieved higher accuracy, demonstrating its robustness and effectiveness in cancer classification. The results indicate that integrating transformer-based architectures can enhance predictive accuracy, making it a promising approach for gene expression analysis. By comparison with recent works, Table 9 indicates the comparison between those papers referenced with the proposed method. One limitation of our study model was that it was trained on specific GEO and TCGA datasets, potentially limiting its direct applicability to unseen cancer types or data modalities. The interpretability of the deep learning model also requires further investigation.

7. CONCLUSION

Gene expression profiling for early cancer diagnosis is a new strategy that is intended to help with the early detection and treatment of several types of cancer. In this research, we proposed a DistilBERT model as a multi-class classifier to classify different types of cancer from a variety of sources, including a cancer dataset. We obtained lung and ovarian cancer from GEO, which provides microarray datasets, and used each one separately. We downloaded (BRCA, KIRC, COAD, LUAD, and PRAD) from TCGA, which offered RNA-Seq datasets, which were then merged to create a substantial dataset for cancer classification. We employed a self-attention mechanism to select important features in the dataset and compare the performance of our proposed method with other models and techniques that are used in ML to classify cancer types. We conclude that our proposed model achieved the highest performance compared to other ML methods and techniques. As a result, our proposed approach can accurately categorize all of the observed positive cancer cases. The suggested model can improve early identification of cancer susceptibility, guiding early intervention decisions and ultimately improving survival rates. The suggested model surpasses others across all datasets, achieving the highest classification accuracy: 97.56% for lung cancer, 100% for ovarian cancer, and 99.504% for the TCGA dataset, which includes five types of cancer. In the future, we plan to boost the quality of gene expression data and use metaheuristic optimization alongside deep learning to take our performance to the next level, and we will explore metaheuristic optimization for feature selection and hyperparameter tuning. We also aim to evaluate the model on broader datasets.

REFERENCES

[1] F. Aldi, F. Hadi, N. A. Rahmi and S. Defit. "Standardscaler's potential in enhancing breast cancer accuracy using machine learning". *Journal of Applied Engineering and Technological Science*, vol. 5, no. 1, pp. 401-413, 2023.

[2] L. Rukhsar, W. H. Bangyal, M. S. Ali Khan, A. A. Ag Ibrahim, K. Nisar and D. B. Rawat. "Analyzing RNA-seq gene expression data using deep learning approaches for cancer classification". *Applied Sciences,* vol. 12, no. 4, p. 1850, 2022.

[3] N. Tabassum, M. A. S. Kamal, M. Akhand and K. Yamada. "Cancer classification from gene expression using ensemble learning with an influential feature selection technique". *BioMedInformatics,* vol. 4, no. 2, pp. 1275-1288, 2024.

[4] M. L. R. AbdElNabi, M. Wajeeh Jasim, H. M. El-Bakry, M. Hamed N. Taha and N. E. M. Khalifa. "Breast and colon cancer classification from gene expression profiles using data mining techniques". *Symmetry*, vol. 12, no. 3, p. 408, 2020.

[5] H. AlShamlan and H. AlMazrua. "Enhancing cancer classification through a hybrid bio-inspired evolutionary algorithm for biomarker gene selection". *Computers, Materials and Continua,* vol. 79, no. 1, pp. 675-694, 2024.

[6] W. Ali and F. Saeed. "Hybrid filter and genetic algorithm-based feature selection for improving cancer classification in high-dimensional microarray data". *Processes,* vol. 11, no. 2, p. 562, 2023.

[7] M. Khalsan, L. R. Machado, E. S. Al-Shamery, S. Ajit, K. Anthony, M. Mu and M. O. Agyeman. "A survey of machine learning approaches applied to gene expression analysis for cancer prediction". *IEEE Access,* vol. 10, pp. 27522-27534, 2022.

[8] R. K. Singh and M. Sivabalakrishnan. "Feature selection of gene expression data for cancer classification: A review". *Procedia Computer Science*, vol. 50, pp. 52-57, 2015.

[9] F. Alharbi and A. Vakanski. "Machine learning methods for cancer classification using gene expression data: A review". *Bioengineering*, vol. 10, no. 2, p. 173, 2023.

[10] S. Gupta, M. K. Gupta, M. Shabaz and A. Sharma. "Deep learning techniques for cancer classification using microarray gene expression data". *Frontiers in Physiology*, vol. 13, p. 952709, 2022.

[11] M. Mohammed, H. Mwambi, I. B. Mboya, M. K. Elbashir and B. Omolo. "A stacking ensemble deep learning approach to cancer type classification based on TCGA data". *Scientific Reports*, vol. 11, no. 1, p. 15626, 2021.

[12] D. Mukhopadhyay, D. D. Phanord, R. J. Dalpatadu, L. P. Gewali and A. K. Singh. "ML classification of cancer types using high dimensional gene expression microarray data". Preprints. 2024.

[13] B. Büyüköz, A. Hürriyetoğlu and A. Özgür. "Analyzing ELMo and DistilBERT on Socio-political News Classification". In: *Proceedings of the Workshop on Automated Extraction of Socio-political Events from News 2020*. Marseille, France, pp. 9-18, 2020.

[14] Y. Wu, Z. Jin, C. Shi, P. Liang and T. Zhan. "Research on the application of deep learning-based BERT model in sentiment analysis". *arXiv preprint arXiv:2403.08217,* 2024.

[15] S. Jamshidi, M. Mohammadi, S. Bagheri, H. E. Najafabadi, A. Rezvanian, M. Gheisari, M. Ghaderzadeh, A. S. Shahabi and Z. Wu. "Effective text classification using BERT, MTM LSTM, and DT". *Data and Knowledge Engineering*, vol. 151, p. 102306, 2024.

[16] Y. Ji, Z. Zhou, H. Liu and R. V. Davuluri. "DNABERT: Pre-trained bidirectional encoder representations from transformers model for DNA-language in genome". *Bioinformatics*, vol. 37, no. 15, pp. 2112-2120, 2021.

[17] E. C. Garrido-Merchan, R. Gozalo-Brizuela and S. Gonzalez-Carvajal. "Comparing BERT against traditional machine learning models in text classification". *Journal of Computational and Cognitive Engineering*, vol. 2, no. 4, pp. 352-356, 2023.

[18] V. Dogra, A. Singh, S. Verma, Kavita, N. Jhanjhi and M. Talib. "Analyzing DistilBERT for Sentiment Classification of Banking Financial News". In: *Intelligent Computing and Innovation on Data Science: Proceedings of ICTIDS 2021*. Springer, Singapore, pp. 501-510, 2021.

[19] S. Sucharita, B. Sahu and T. Swarnkar. "Efficient Gene expression data analysis using ES-DBN for microarray cancer data classification". *EAI Endorsed Transactions on Pervasive Health and Technology*, vol. 10, pp. 1-12, 2024.

[20] H. Hijazi and C. Chan. "A classification framework applied to cancer gene expression profiles". *Journal of Healthcare Engineering*, vol. 4, no. 2, pp. 255-283, 2013.

[21] T. Thakur, I. Batra, A. Malik, D. Ghimire, S. H. Kim and A. S. Hosen. "RNN-CNN based cancer prediction model for gene expression". *IEEE Access*, vol. 11, pp. 131024-131044, 2023.

[22] I. Guyon, J. Weston, S. Barnhill and V. Vapnik. "Gene selection for cancer classification using support vector machines". *Machine Learning*, vol. 46, pp. 389-422, 2002.

[23] Y. Wei, M. Gao, J. Xiao, C. Liu, Y. Tian and Y. He. "Research and implementation of cancer gene data classification based on deep learning". *Journal of Software Engineering and Applications*, vol. 16, no. 6, pp. 155-169, 2023.

[24] Y. Li, K. Kang, J. M. Krahn, N. Croutwater, K. Lee, D. M. Umbach and L. Li. "A comprehensive genomic pan-cancer classification using the cancer genome atlas gene expression data". *BMC Genomics*, vol. 18, p. 508, 2017.

[25] P. García-Díaz, I. Sánchez-Berriel, J. A. Martínez-Rojas and A. M. Diez-Pascual. "Unsupervised feature selection algorithm for multiclass cancer classification of gene expression RNA-Seq data". *Genomics*, vol. 112, no. 2, pp. 1916-1925, 2020.

[26] L. P. Chen. "Classification and prediction for multi-cancer data with ultrahigh-dimensional gene expressions". *PLos One*, vol. 17, no. 9, p. e0274440, 2022.

[27] A. Das, N. Neelima, K. Deepa and T. Özer. "Gene selection based cancer classification with adaptive optimization using deep learning architecture". *IEEE Access*, vol. 12, pp. 62234–62255, 2024.

[28] A. Yaqoob, N. K. Verma and R. M. Aziz. "Optimizing gene selection and cancer classification with hybrid sine cosine and cuckoo search algorithm". *Journal of Medical Systems*, vol. 48, no. 1, p. 10, 2024.

[29] S. Tarek, R. Abd Elwahab and M. Shoman. "Gene expression based cancer classification". *Egyptian Informatics Journal*, vol. 18, no. 3, pp. 151-159, 2017.

[30] S. Aburass, O. Dorgha and J. Al Shaqsi. "A hybrid machine learning model for classifying gene mutations in cancer using LSTM, BiLSTM, CNN, GRU, and GloVe". *Systems and Soft Computing*, vol. 6, p. 200110, 2024.

[31] J. A. Martínez Logreira. "*Machine learning-based cancer classification using gene expression data*", (Master's thesis). Universidad de los Andes, Bogotá, Colombia, 2020.

[32] F. Neutatz, B. Chen, Z. Abedjan and E. Wu. "From cleaning before ML to cleaning For ML". *IEEE Data Engineering Bulletin*, vol. 44, no. 1, pp. 24-41, 2021.

[33] L. Huang, J. Qin, Y. Zhou, F. Zhu, L. Liu and L. Shao. "Normalization

techniques in training dnns: Methodology, analysis and application". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 45, no. 8, pp. 10173-10196, 2023.

[34] J. Sun and Y. Xia. "Pretreating and normalizing metabolomics data for statistical analysis". *Genes and Diseases*, vol. 11, no. 3, p. 100979, 2024.

[35] R. Dang and W. Yu. "Standard deviation effect of average structure descriptor on grain boundary energy prediction". *Materials*, vol. 16, no. 3, p. 1197, 2023.

[36] Z. Huo, G. Du, F. Luo, Y. Qiao and J. Luo. "D-MSCD: Mean-standard deviation curve descriptor based on deep learning". *IEEE Access*, vol. 8, pp. 204509-204517, 2020.

[37] R. Pramanik, B. Banerjee and R. Sarkar. "MSENet: Mean and standard deviation based ensemble network for cervical cancer detection". *Engineering Applications of Artificial Intelligence*, vol. 123, p. 106336, 2023.

[38] Y. Chen, X. Kou, J. Bai and Y. Tong. "Improving bert with self-supervised attention". *IEEE Access*, vol. 9, pp. 144129-144139, 2021.

[39] B. Cui, Y. Li, M. Chen and Z. Zhang. "Fine-tune BERT with Sparse Self-attention Mechanism". In: *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing* (*EMNLP-IJCNLP*). Association for Computational Linguistics, Hong Kong, China, pp. 3548-3553, 2019.

[40] B. Ghojogh and A. Ghodsi, "*Attention mechanism, transformers, BERT, and GPT: Tutorial and survey*," [Preprint] 2020.

[41] Y. Hao, L. Dong, F. Wei and K. Xu. "Self-attention attribution: Interpreting information interactions inside transformer". *Proceedings of the AAAI Conference on Artificial Intelligence*, 2021, vol. 35, no. 14, pp. 12963-12971.

[42] J. Shobana and M. Murali. "An improved self attention mechanism based on optimized BERT-BiLSTM model for accurate polarity prediction". *The Computer Journal*, vol. 66, no. 5, pp. 1279-1294, 2023.

ORIGINAL RESEARCH ARTICLE UHD JOURNAL OF SCIENCE AND TECHNOLOGY

Improving Cardiovascular Disease Prediction through Stratified Machine Learning Models and Combined Datasets

Tara Yousif Mawlood[1], Alla Ahmad Hassan[2], Rebwar Khalid Muhammed[3], Aso M. Aladdin[4*], Tarik A. Rashid[5], Bryar A. Hassan[5]

[1]Department of IT, Computer Science Institute, Sulaimani Polytechnic University, Sulaymaniyah, Iraq, [2]Department of Database, Computer Science Institute, Sulaimani Polytechnic University, Sulaymaniyah, Iraq, [3]Department of Network, Computer Science Institute, Sulaimani Polytechnic University, Sulaymaniyah, Iraq, [4]Department of Computer Science, College of Science, Charmo University, Sulaymaniyah, Iraq, [5]Department of Computer Science and Engineering, School of Science and Engineering, University of Kurdistan Hewler, Erbil, Iraq

ABSTRACT

The global rise in cardiovascular disease (CVD) cases underscores the critical need for accurate and early diagnostic solutions. This study introduces a robust machine learning (ML) framework for predicting CVD risk by integrating two large, feature-identical datasets containing clinical and biological indicators along with patient history. Seven classification algorithms – logistic regression, random forest (RF), support vector machine (SVM), Gaussian naive Bayes (GNB), gradient boosting (GB), K-nearest neighbors, and decision tree (DT) – were employed. A stratified sampling strategy was used to ensure balanced class distribution, and model performance was further validated using k-fold cross-validation to enhance robustness and generalizability. The datasets, sourced from the UCI repository, were pre-processed and evaluated using metrics such as accuracy, precision, F1-score, log loss, and error rate, with performance further assessed using confusion matrices. Results revealed that ensemble models, particularly RF and DT, achieved optimal performance with 100% accuracy, while stratification significantly improved the outcomes of SVM, GNB, and GB. The integration of datasets, stratified sampling, and k-fold validation effectively enhanced model reliability while minimizing overfitting. These findings highlight the potential of ML to support early CVD diagnosis and lay the groundwork for future research on hybrid models and real-world clinical applications.

Index Terms: Cardiovascular Disease, Gradient Boosting, Heart Disease, K-Nearest Neighbors, Logistic Regression, Naive Bayes, Support Vector Machine.

1. INTRODUCTION

The cardiovascular system, also known as the circulatory system, is considered one of the most vital systems in the human body, along with the liver, lungs, and other essential organs [1], [2]. Cardiovascular disease (CVD) has become one of the most prevalent illnesses in nations worldwide today. It is often caused by low blood and oxygen levels in the circulatory system as well as blood vessel stenosis [3], [4]. Medical centers today have access to numerous datasets specifically focused on heart disease diagnoses. The evaluation of diseases and recognizing objects both make significant utilization of machine learning (ML) techniques [5]. Using ML algorithms to diagnose diseases, enormous quantities of medical data are converted into information

Access this article online

DOI:10.21928/uhdjst.v9n1y2025.pp149-168 E-ISSN: 2521-4217 P-ISSN: 2521-4209

Corresponding author's e-mail: Aso M. Aladdin, Department of Computer Science, College of Science, Charmo University, Sulaymaniyah, Iraq. E-mail: aso.aladdin@chu.edu.iq

Received: 12-02-2025 Accepted: 11-05-2025 Published: 03-06-2025

that can improve predicting and decision-making. To help healthcare providers by improving the precision and accuracy of making decisions over disease detection and diagnosis, ML research has gained importance in healthcare. A pair of its goals is the development of machine-based evaluation systems and disease prediction [6], [7]. A number of symptoms are able to identify CVD: Hypertension, chest pain, high blood pressure, cardiac arrest, etc. Many CVD types are present, each having a different variety of symptoms. Such as: (1) Chest pain, dyspnea, and throat pain are symptoms of cardiac disease in the blood vessels. (2) CVD caused by irregular heartbeats: discomfort, a slowing heartbeat, chest pain, etc. The most typical symptoms are discomfort, shortness of breath, chest pain, etc. The most typical symptoms include fainting, shortness of breath, and chest pain. CVD can be led on by pre-mature births, diabetes, high blood pressure, cigarette smoking, drug usage, and drinking alcohol. A fever, exhaustion, dry cough, and skin rashes are signs that the infection has occasionally spread to the inner membranes of the heart. Bacteria, viruses, and parasites are the causes of heart infections. Heart disease kinds include the following: Angina pectoris, congenital heart disease, Cardiac failure, cardiac illness, high blood pressure, plaque in the arteries, and a slower beat of the heart. Many automated techniques are available nowadays, including deep neural networks, algorithmic learning, and data mining [18]. Heart disease risk is commonly predicted based on various factors such as insulin resistance, CVD, high bad cholesterol levels, age, gender, smoking or drinking habits, obesity, heart rate, and chest pain, as demonstrated in numerous previous studies on different cases [9]. Leveraging medical information gathered from real-world cases, technology has become increasingly effective in predicting heart disease. With advancements in ML, a core component of artificial intelligence, these technologies have reached new heights. ML provides powerful tools for medical diagnosis and disease prediction, significantly enhancing the quality of healthcare services. By analyzing real-life medical data, these systems can more accurately detect whether an individual is at risk for heart disease. Addressing this issue requires robust accuracy in ensuring data security and maintaining confidentiality between patients and physicians. This can be achieved through the implementation of well-established security algorithms [10]. ML technology, considered a component of artificial intelligence, represents the highest point of technological advancement. Due to its ability to provide medical diagnostic tools for disease prediction, ML plays a significant role in enhancing the quality of health services [11]. Despite this, many metaheuristic algorithms have been employed to classify data and optimize this problem through evolutionary evaluation [12], [13], [14]. Furthermore, the supervised ML approach includes an algorithm for classification that can be utilized for prediction [15].

The aim of this study is to determine if cardiac questions can be identified based on a patient's medical factors, such as age, gender, and chest pain. For this purpose, patient characteristics and medical selected data datasets as specified in the future. By analyzing this dataset, the goal is to determine whether a patient has a cardiac issue.

The novelty of this study lies in its innovative approach to enhancing CVD prediction by combining two datasets (UCI and HD) and applying stratified ML techniques. A key aspect of this novelty is the integration of two large datasets with identical feature variables, creating a more diverse and comprehensive dataset that improves model generalization compared to using a single source. In addition, the study employs stratified data splitting to maintain balanced class distributions, which is particularly beneficial for imbalanced datasets. This technique not only reduces model bias but also significantly improves the prediction accuracy of algorithms such as logistic regression (LR), random forest (RF), support vector machine (SVM), Gaussian Naive Bayes (GNB), gradient boosting (GB), K-nearest neighbors (KNN), decision tree (DT). The study further distinguishes itself through a thorough comparative performance evaluation of multiple ML algorithms – including LR, RF, SVM, GNB, GB, KNN, and DT – assessing their effectiveness with and without stratification. Notably, the research provides a quantitative analysis of how stratification positively impacts model accuracy, precision, and F1-score, underscoring its critical role in healthcare-related ML tasks. While the study leverages widely used ML methods, its uniqueness is evident in the dataset integration, the strategic use of stratification, and the detailed assessment of its effects on model performance. To further enhance its contributions, future work could explore advanced techniques such as deep learning (DL), hybrid models, or clinical validation using real-world healthcare data. This study's primary contribution lies in training health-related features using seven methods:

- Evaluate the performance of various ML algorithms, including LR, RF, GNB, GB, KNN, SVM, and DT, in predicting CVD.
- The combined dataset showed enhanced accuracy by using the stratify parameter, which ensured balanced training and improved model performance.
- This strategy helps medical professionals assess patient risk, showcasing ML's potential to enhance diagnostic accuracy and improve patient outcomes.

- The use of stratified techniques greatly enhanced the accuracy of SVM, GNB, and GB models, emphasizing the value of balanced data distribution during training.
- This study evaluates the performance of various classification algorithms on three distinct datasets: UCI, HD, and the combined UCI-HD dataset.
- The results highlight the effectiveness of RF and DT algorithms, demonstrating their robustness across all datasets, with RF achieving perfect accuracy on the combined dataset. In addition, the performance of other algorithms, such as KNN and GNB, provides valuable insights into their strengths, particularly on the HD dataset.
- This comparative analysis supports the selection of the most suitable algorithm for classification tasks across different datasets.

The rest of the paper is organized as follows: Section 2 delivers a background review, followed by Section 3, which outlines the methods and materials used. Section 4 presents the results and analysis, while Section 5 discusses the performance of the algorithms based on the adjustments. Finally, the conclusion and future work are presented in the past section. Procedure for Paper Submission.

2. BACKGROUND REVIEW

Many studies involving animals or humans, and other studies that require ethical approval, must list the authority that provided approval and the corresponding ethical approval code. Many studies are being carried out having the goal of using algorithms based on ML to identify cardiac disease. The study employed various ML methods to create a prediction model for categorizing cardiac diseases. The following part discusses a few of the earlier studies on predicting the likelihood of heart disease.

The DT, RF, and Naïve Bayes (NB) methods are applied to the Cleveland heart disease dataset by Gavhane *et al.* [16]. The study dataset was used to evaluate the accuracy of predictions of the approaches, and the RF strategy performed better than the DT and NB procedures. Ambekar and Phalnikar [17] use a heart disease dataset to compare the prediction power of many ML methods, such as GNB, LR, RF, and KNN. In terms of prediction accuracy, LR performed better than all other approaches according to the two different findings. The comparison of three ML algorithms, DT, RF, and multi-layer perception, using the Wisconsin Heart Disease Data Repository is carried out by Jothi *et al.* [18]. The methods are compared for accuracy in predicting cardiac illness, and the findings indicate that multi-layer perception and neural networks perform better in this regard. The Wisconsin Heart Disease Data Repository forecasts heart disease using NB, as described by Segie *et al.* [19]. It can achieve 87% overall prediction accuracy. In artificial ML and forecasting systems, the NB model performs better than the other models in terms of performance and the ability to accurately forecast cardiovascular illness, with an adequate forecasting accuracy of eighty-seven. The study conducted by Kajal and Nishika [20] focused on different methods of classification used to predict an individual's danger degree based on blood pressure, cholesterol levels, cardiac rate, age, gender, and other characteristics.

Data mining methods, including NB, KNN, DT, and Neural Network, are used to increase the accuracy of the hazard level. The kernel nearest neighbor and ID3 algorithms were used to determine the risk rate of heart disease. The accuracy rating for various amounts of attributes was also provided. Babu *et al.* [21], conducted research on a range of educational apps that support the identification of many cardiac conditions. A selection of methods, including data mining, SVMs, computationally intelligent classifiers, and hidden markov's models, were employed. Since treating heart disease is extremely costly and out of reach for the average person, these kinds of cutting-edge technologies are created to address this issue. The beginning predictions can also benefit from these strategies. It modifies daily routines somewhat to prevent more suffering. As a result, the author draws the conclusion that the anticipated strategy is highly helpful and offers several advantages. Kannan and Vasanthi [22] employed a variety of ML algorithms, including LR, RF, SVM, and stochastic gradient boosting, to identify potential cardiovascular illnesses. The equation predicts that LR has the most accuracy, coming in at 86.5%.

In addition, Raza [23] employed LR, NB, a multilayer perceptron, and a combined learning model to classify heart diseases. The result shows that combined learning has improved the prediction performance of cardiac disease when compared to other approaches. For example, Z-Alizadesh Sani and the Cleveland heart disorder dataset were two separate datasets utilized by Sapra *et al.* [24] to diagnose cardiac illnesses. These datasets were previously analyzed using six different ML techniques: LR, DL, DT, RF, SVM, and collaborative learning gradient boosted tree. When compared to other techniques, gradient boosted tree strategy yielded the highest accuracy, at 84%. Because CVD

has one of the highest death rates worldwide, projection is an essential part of medicine. Several techniques involving ML are being used to effectively forecast cardiac disease. Research has demonstrated that technologies such as KNN, RF, DT, and SVMs can reliably forecast CVD with excellent accuracy rates of 79%–94% [25]. According to this research [26], NB (83.60%), KNN (90.16%), LR (86.88%), RF (96.72%), extreme gradient boost (95.08%), and DT (77.049%) were the ML techniques that were employed. Remarkably, the method known as RF performed more effectively than the remaining algorithms, with a maximum accuracy rate of 96.72%. For further clarification on the methods used, Table 1 includes a summary of the approaches, along with a comparison to previous works that have addressed similar topics.

3. MATERIALS AND METHODS

3.1. Dataset

The UCI Heart Disease dataset is a widely used dataset for predicting CVD, containing various attributes related to patient health, such as age, cholesterol levels, and electrocardiographic results. The investigation utilized two datasets for analysis, as summarized in Table 2. The first dataset, UCI, contains samples categorized based on the presence or absence of CVD. Similarly, the second dataset, HD, provides additional samples with the same classification criteria.

The Combined dataset, created by merging UCI and HD, offers a comprehensive view of CVD distribution. All

TABLE 1: Relevant bibliography

Title of research	The methods	Different types of data used	Year	References
Utilizing Artificial Intelligence to Predict Cardiovascular Disease	RF, DT, and NB	Cleveland Cardiovascular Disease Data Set	2018	[16]
Making use of convolutional neural networks to predict the probability of diseases	GNB, RF, KNN, and L R	Cardiovascular Disease Data Set	2018	[17]
Genetic algorithm-based feature determination technique in the healthcare dataset	RF, DT, and multi-layer neural networks for perception	Wisconsin Cardiovascular Research Center	2019	[18]
Support Vector Techniques for Machine-Based Detection of Cardiovascular Disease	NB	Wisconsin Cardiovascular Research Center	2019	[19]
Cardiovascular Disease Detection Utilizing Data Mining Methods	Neural Network, KNN, DT, and NB	Features: Sexuality, Age, Blood Pressure, Cholesterol Levels, and Heart Rate	2016	[20]
Diagnosing Cardiovascular Disease Through Data Mining Method	Data mining, computerized smart classifiers, SVM, and concealing Markov chain	Cardiovascular Disease Data Set	2017	[21]
ROC curve-based machine learning techniques for cardiovascular disease diagnosis and prognosis	Probabilistic Gradient Boosting, RF, SVM, and LR	Cardiovascular Disease Data Set	2019	[22]
Improving the identification of cardiovascular disease accuracy using a majority vote and group learning	LR, NB, Multilayer Perceptron, Ensemble Learning	Cardiovascular Disease Data Set	2019	[23]
Using a combined technique, a smart approach for detecting CAD	Gradient Boosted Tree, RF, SVM, DT, DL, and LR	Cleveland Cardiovascular Disease and Z- Alizadesh Sani Datasets	2021	[24]
Using Machine Learning Techniques to Track Cardiovascular Issues in Cardiovascular Disease	SVM, RF, DT, and KNN	Cardiovascular Disease Data Set	2023	[25]
An Organized Analysis of Machine Learning Algorithms for Heart Failure Prediction	KNN, LR, and RF	Cardiovascular Disease Data Set	2022	[25]

LR: Logistic regression, RF: Random forest, SVM: Support vector machine, GNB: Gaussian Naive Bayes, KNN: K-nearest neighbors, DT: Decision tree, DL: deep learning

TABLE 2: Cardiovascular disease dataset distribution and sources

Dataset	Feature	Samples	Negative samples (%)	Positive samples (%)	Source
UCI	14	1025	499 (48.68)	526 (51.32)	https://www.kaggle.com/datasets/johnsmith88/heart-disease-dataset/data
HD	14	303	138 (45.5)	165 (54.5)	https://www.kaggle.com/code/mragpavank/heart-disease-uci/notebook
UCI-HD	14	1328	637 (48.0)	691 (52.0)	Includes UCI -HD

datasets are complete, with no missing feature values. Fig. 1 illustrates the distribution of CVD across the samples. For model development, 80% of the combined dataset was allocated for training, and 20% for testing.

3.2. Data Pre-Processing

This study focuses on pre-processing the UCI and HD (heart disease) datasets before developing a predictive model using ML. These datasets have undergone extensive cleaning and pre-processing, making them easier to use and requiring minimal effort for data preparation. In addition, they are well-documented and frequently cited in scientific research. In both datasets, the target attribute is an integer indicating the presence of heart disease in a patient. A value of 0 signifies no heart disease, while a value of 1 indicates its presence. The attribute sex represents gender, with two classes: 1 for males and 0 for females. The attribute cp (chest pain type) includes four classes, while fbs (fasting blood sugar) has two classes. Similarly, restecg (resting electrocardiogram) consists of three classes, and exang (exercise-induced angina) has two classes. The attribute slope (ST slope) comprises three classes. Four additional attributes – trestbps (resting blood pressure), chol (cholesterol level), age, and oldpeak – are treated as numerical values. The data pre-processing process involves multiple steps, from data loading to splitting for training and testing. These steps are detailed in Table 3.

3.3. Selection Algorithms

There are several algorithms commonly used for the classification of cardiac disease, including LR, RF, SVM, GNB, GB, KNN, and DT.

3.3.1. LR algorithm

It consists of a classification process that makes use of only one class-based classifier and a single multinomial LR approach. When using a specific technique, a LR analysis

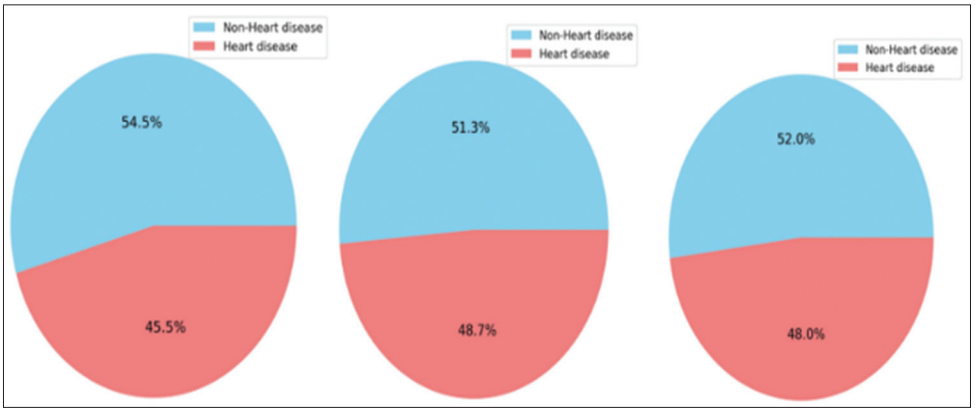


Fig. 1. Patient with cardiovascular disease compared to non-heart disorder patients.

TABLE 3: Heart dataset features

Attributes	Information
Sex	Gender of participants
Cp	Kind of chest pain. There are four parts to this feature: unusual angina, usual angina, and both. Pain that is not anginal and asymptomatic
Trestbps	Blood pressure of the patient during a time of rest or inactivity
Chol	Level of cholesterol
Fbs	The level of blood sugar has an accurate level if it is more than 120 mg/dL and an incorrect value if it is below 120 mg/dL.
Restecg	Electrocardiogram outcomes acquired as a patient is at rest, often known as resting electrocardiogram results. Following the Estes criteria, a result of 0 denotes normalcy, an amount of 1 suggests aberrant ST-T waves, & a result of 2 suggests a certain risk of hypertrophy of the left ventricle.
Thalach	Maximal heartbeat
Exang	The patient's pain whereas physical activity. Value, true if the answer is "yes" and false if the word "no"
Oldpeak	The reduction in ST brought on by activity.
Slope	Slope during exercise at maximum ST. The slope of it can be classified as upsloping, round, or downsloping.
Ca	The amount of vessels is identified by coloring.
Thal	Test for thalassemia has a total of three numbers: changeable error, immovable defect, and regular.
Target	Cardiac illness (1) is a type of label. Non-cardiac illness (0)
Thalach	Maximal heartbeat

usually shows where the group's borders are and how far removed the category's probabilities are from them. This approaches the extremes (0 and 1) more quickly depending on the data collection. LR is elevated beyond the level of a basic classification by these conditional arguments. It might be applied in a different method more offers more accurate and thorough estimates, although they are not certainties. LR is an estimating technique that resembles ordinary least squares regression. However, employing LR for prediction results in a single response [27]. LR has become one of the increasingly popular techniques for intermittent information analysis and statistical applications. The LR approach makes use of the linear interpolation method [28]. The sigmoid function, given by the equation (1)

$$sig(t) = \frac{1}{1+e^{-t}} \tag{1}$$

3.3.2. RF algorithm

Among the most effective learning techniques is RF. Humanities researchers are able to benefit from algorithmic advancements if they are allowed access to an application of the method. Since models that utilize trees are the foundation of the RF technique, let's start by talking about them. With a tree-based approach, the collection of data is repeatedly divided into two separate categories according to a pre-defined parameter or another is satisfied. The referred to as nodes of leaves, or leaves are located at the base of the DT. Comparing to DTs, the RF algorithm calculates the error rate more accurately. Particularly, it can be demonstrated analytically that the error rate constantly converges as the number of trees rises [29]. At the preparation phase, the out-of-bag (OOB) error approximates the variance of the RF mistake. A distinct bootstrapping sample serves as the foundation for every tree. Approximately one-third of the observations are randomly excluded from each bootstrapping dataset. An OOB example of this is the collection of each of this excluded information for an individual tree. When choosing an algorithm and fine-tuning factors, determining which ones will result in a small OOB variance is frequently crucial. Keep in mind that the size of the group of predictor factors is essential for regulating the trees' ultimate depth in the RF method. As a result, it is a parameter that must be adjusted when choosing a model.

3.3.3. SVM algorithm

The most commonly implemented supervised ML technique, SVM, is utilized for classification as well as regression. However, this approach is mostly examined for ML issues with classification. To swiftly place the newly acquired information in the right group, the strategy known as SVM aims to construct the most effective border, the line of sight that may divide the space with n dimensions into classes. This optimal selection of boundaries is referred to as a "hyperplane" [30]. SVM selects the extreme vectors that help to create the hyperspace. The highest and lowest vectors are collectively referred to as vectors of support, and the method that employs a disproportionate number of verticals is called SVM. The SVM image below classifies two distinct groups using hyperplanes or borders of selections.

3.3.4. KNN algorithm

The KNN algorithm, the most basic categorizing technique, is based on supervised learning techniques. The KNN technique can be applied to reversion, although it is primarily used for classification [22]. A new data point is categorized using the KNN algorithm according to how well the information it provides matches the existing data. How fresh data can be promptly classified by the method known as KNN when it falls into an appropriate class.

3.3.5. GNB algorithm

The GNB, which comes through the Bayesian theory, GNB provides constant data that are drawn through the Gaussian typical distribution. The idea that its elements are autonomous is the foundation of the GNB. This sorting algorithm is regarded as being one of the most straightforward and practical methods. For categorization using supervised ML, which is predicated on the concept that the information is regularly generated. GNB is demonstrated [31]. If every chosen feature contributes equally and independently, the Bayes theorem depends on multiplying the probability and previous by the evidence presented. Since they are independent of one another, it is assumed that the significance of each attribute has an equal impact on the result. The likelihood of a specific event because something else has previously occurred is known to be called the GNB using the formula (2).

$$(A|B) = \frac{P(B|A)\ P(A)}{P(B)} \tag{2}$$

This formula determines the following probability, or the likelihood that A will occur, considering that B will occur, and the likelihood that B will occur, provided that A will occur. The probabilities of A Occurring P(A) split by proof, or the possibility of B occurring P(B), is P(B|A), which is the probability multiply by the previous event.

3.3.6. GB

In terms of accuracy, the GB provides a cutting edge, particularly for supervised training tasks on structured datasets. Freund and Schapire created AdaBoost, the initial boosting method, in the year 1997. Combining various models of ML is the aim of collaborative learning with the objective of enhancing the accuracy of predictions. The precise concept behind boosting is that you should start with an algorithm that is referred to as a poor learner, meaning it is just marginally higher quality than chance guesswork. Over time, the algorithm gets better by fixing the mistakes made by the prior model at each stage. GB begins with a poor learner, usually a DT, and continuously refines this starting learner by accounting for the inaccuracy of the preceding models at each stage. GB has become considered among the most successful algorithmic learning strategies available. The most common form of boosting progressively advertises a single DT as time by using trees of choices. Higher accuracy is the result of the model's effectiveness being gradually improved by this sequential change. There are many kinds of GB available. The gradual boosting version used in this investigation relies on [32].

3.3.7. DT algorithm

The DT Algorithm is a supervised learning method for classification and regression. It splits data into subsets based on attribute values, forming a tree-like structure of decisions and outcomes. Metrics, such as entropy and information gain [33]. The entropy value is calculated using the formula in Equation 3, when S is representing a specific set, A is one of the attributes, the number of divisions for attributes is n, (|Si|) is the quantity of instance in partition (i), finally |S| is the sum of the instance in (S).

$$\text{Gain}(S,\ A) = \sum_{i=1}^{n} \frac{|S1|}{S}\ \text{entropy}(\text{Si}) \tag{3}$$

3.4. Proposed Method

The proposed method, illustrated in Fig. 2, flowchart for the Proposed System, outlines a comprehensive approach to heart disease prediction using ML. The process begins with

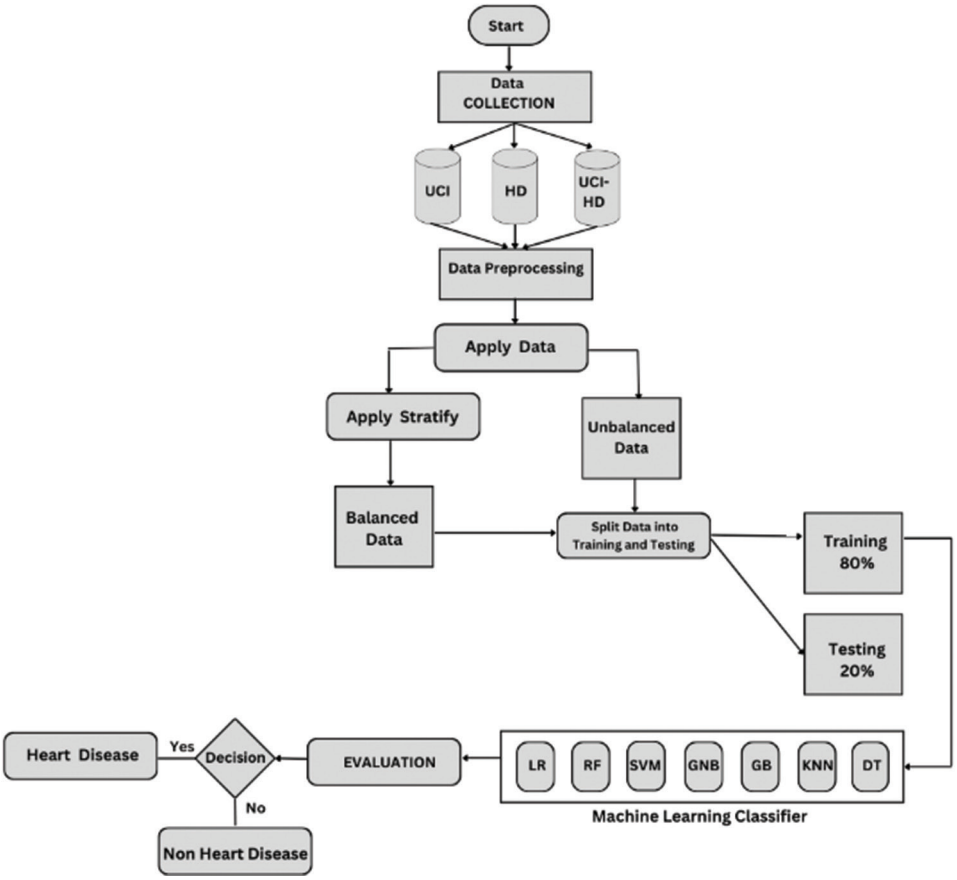


Fig. 2. Flowchart for the proposed system.

TABLE 4: Performance evaluation of machine learning algorithms on the UCI dataset

Algorithm	Accuracy	Precision	F1-score	Log loss	Error rate
LR	79.51219	0.80234	0.7937867	0.45313	0.2048780
RF	98.53658	0.98578	0.9853637	0.06658	0.0146341
SVM	88.78048	0.89226	0.8874512	0.25314	0.1121951
GNB	80.00000	0.81050	0.7981741	0.66268	0.2000000
GB	93.17073	0.93235	0.9316748	0.19782	0.0682926
KNN	83.41463	0.83869	0.8335281	0.25111	0.1658536
DT	98.5365	0.98578	0.985363	0.5274	0.014634

LR: Logistic regression, RF: Random forest, SVM: Support vector machine, GNB: Gaussian Naive Bayes, GB: Gradient boosting, KNN: K-nearest neighbors, DT: Decision tree

data collection from three sources: UCI, HD, and UCI-HD datasets, ensuring a diverse and robust dataset. This data undergoes pre-processing to clean and normalize it, addressing issues, such as missing values, noise, and inconsistencies to prepare it for analysis. The next step involves applying the data. For balanced datasets, a stratification process is applied to ensure even class distribution, reducing bias. For unbalanced datasets, the data are split into training (80%) and testing (20%) subsets, enabling effective model evaluation. The processed data are fed into a ML classification pipeline consisting of algorithms, such as LR, RF, SVM, GNB, GB, KNN, and DT. These classifiers are trained on the dataset to build predictive models. Finally, the system undergoes an evaluation phase, classifying outcomes into two categories: Heart Disease and Non-Heart Disease, based on the decision boundary of the classifiers. This robust framework ensures accurate and fair predictions, with balanced data enhancing model reliability and unbalanced data reflecting real-world scenarios.

3.4.1. The importance of stratified sampling in dataset splitting

Table 4 shows that stratify = y is used in train_test_split, it ensures that the class distribution in the target labels (y) is maintained across both training and test sets. This is particularly useful for imbalanced datasets, where some classes may have fewer samples than others. The function groups the data by unique classes in y and splits each class proportionally into training and testing subsets. This process ensures that the relative frequency of each class in y remains consistent. If stratify is not specified, the split is purely random and may lead to class imbalances in the subsets.

The data processing in this study employs a stratified splitting method to ensure that the class distribution of the target labels is preserved across both training and testing subsets. The dataset is split into two parts: A training set and a testing set, with the proportion for the testing set specified by the test_size parameter (20%). The random_state seed ensures that the data split is reproducible. If the stratify option is enabled, the dataset is grouped by unique target label classes, and each group is split into training and testing sets while maintaining the original class proportions. These stratified splits are then combined into the final training and testing sets. If stratification is not applied, the function will randomly shuffle and split the data based on the test_size ratio. This stratified methodology ensures that the subsets are representative of the overall dataset, preserving the balance of target labels in both training and testing sets, which is critical for training accurate ML models. Fig. 3 illustrates the stratified dataset splitting for training and testing.

```
# Inputs:
# x: Feature dataset (0 to N)
# y: Target labels indicating Heart Disease Status (0: No, 1: Yes)
# test_size: Proportion of the dataset to be used for testing (e.g., 0.20 for 20%)
# random_state: Seed for reproducibility (e.g., 42)
# stratify: Use target labels to maintain class distribution (typically set to 'y')

# Output:
# x_train, x_test: Feature subsets for training and testing
# y_train, y_test: Corresponding labels

# Description:
# Function to split the dataset into training and testing sets.
# If stratification is applied, it ensures both sets maintain the original class distribution.

FUNCTION train_test_split(x, y, test_size, random_state, stratify):
  SET seed = random_state
  SET test_size_ratio = test_size

  IF stratify is provided:
    GROUP the data in x and y by the unique classes in 'stratify'
    SPLIT each class group into training and testing subsets with the same class proportions
    COMBINE all stratified subsets into final training and testing sets
  ELSE:
    RANDOMLY shuffle x and y using the seed
    SPLIT the shuffled data into training and testing sets based on test_size_ratio

  RETURN x_train, x_test, y_train, y_test
END
```

Fig. 3. Stratified dataset splitting for training and testing.

4. RESULTS

According to the methodology of this study, the datasets have been compared based on the comparative parameters for each algorithm described in subsection one. Notably, the

results for each dataset are presented and the modifications evaluated are illustrated in subsection two.

4.1. Comparative Parameters

The performance of the proposed model was evaluated using the UCI, HD, and Combined (UCI-HD) datasets, which provide comprehensive information on CVD. After dividing the data, the model was trained and tested using algorithms such as LR, RF, GNB, GB, SVM, DT, and KNN. The algorithm with the highest efficiency was identified by analyzing performance metrics, including accuracy, precision, F1-score, and logarithmic loss. Accuracy, which measures the percentage of correctly classified samples, was calculated using the formula derived from the confusion matrix. This formula, referred to as equation 4, quantifies the model's ability to classify CVD cases accurately across different datasets.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (4)$$

Precision evaluates the accuracy of a classifier by comparing the number of true positives (TP) in the actual data to the number of predicted TP. This measure of accuracy is essential for assessing the performance of the proposed method, as calculated mathematically according to equation 5.

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

F1-Score: The F1-score is a statistical metric used to evaluate the performance of a classification model. It provides a balanced assessment by calculating the harmonic mean of precision and recall. This single metric considers both recall and precision, offering a comprehensive evaluation of model performance. The F1-score is calculated as the harmonic mean of these two values, with the formula for the F-measure represented in equation 6.

$$F - measure = \frac{2 * (Precision * Recall)}{Precision + Recall} \quad (6)$$

The performance of a classifier can be effectively represented and evaluated using a confusion matrix. TP represents the number of individuals correctly identified as having the disease. True Negatives represent the number of individuals correctly identified as not having the disease. False Positives (FP) represent the number of healthy individuals who are incorrectly diagnosed with the disease. False Negatives (FN) occur when individuals with the disease are incorrectly classified as healthy.

4.2. Evaluative Results

The datasets are utilized to identify the most effective model for predicting cardiovascular disease in patients. This investigation is based on established algorithms commonly used in healthcare predictions. The nominated model demonstrated particularly strong performance in specific cases, making it a valuable tool for professionals in diagnosing this condition. All issues, comparisons, and outcomes are thoroughly illustrated in the subsequent tables.

Table 4 shows the performance evaluation of ML algorithms on the UCI dataset reveals notable differences in accuracy, precision, F1-score, log loss, and error rate. RF and DT achieved the highest accuracy (98.54%) and precision (0.9858), indicating superior performance. GB followed with 93.17% accuracy and balanced metrics, while SVM demonstrated strong results with an 88.78% accuracy. LR and GNB showed moderate performance, with accuracy values of 79.51% and 80%, respectively. KNN achieved 83.41% accuracy but slightly higher error rates compared to top-performing models. RF and DT stand out as optimal models for the dataset due to their low error rates (1.46%) and minimal log loss. These findings highlight RF and DT as robust classifiers for this dataset.

A comparison of error rates among seven ML algorithms, such as LR, DT, GNB, KNN, GB, RF, and SVM, is presented in Fig. 4. Among these, RF demonstrated the lowest error rate, showcasing its effectiveness and robustness for this dataset. GB and SVM also achieved strong results, with error rates slightly higher than that of RF. LR and GNB showed moderate performance, whereas KNN had a higher error rate. While the DT performed competitively, it was outperformed by the ensemble methods. Overall, ensemble

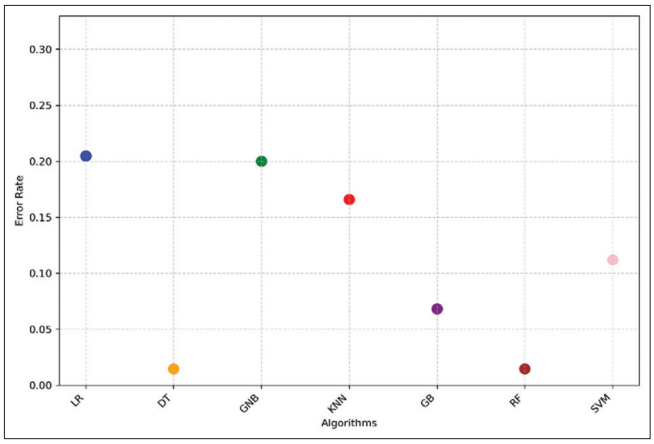


Fig. 4. Comparison of error rates for machine learning algorithms on UCI dataset.

TABLE 5: Performance metrics of classification algorithms on the UCI dataset using stratified sampling

Algorithm	Accuracy	Precision	F1-score	Log loss	Error rate
LR	80.97561	0.822476	0.807244	0.348282	0.190244
RF	100.00000	1.00000	1.00000	0.05885	0.00000
SVM	92.68293	0.927144	0.926787	0.178714	0.073171
GNB	82.92683	0.831469	0.828754	0.506077	0.170732
GB	97.56098	0.975649	0.975606	0.156275	0.02439
KNN	86.34146	0.863618	0.863434	0.216995	0.136585
DT	98.53659	0.985792	0.985368	0.527468	0.014634

LR: Logistic regression, RF: Random forest, SVM: Support vector machine, GNB: Gaussian Naive Bayes, GB: Gradient boosting, KNN: K-nearest neighbors, DT: Decision tree

models, especially RF, proved to be the most effective for this dataset.

To modify these algorithms as stated, a stratified procedure was employed. Table 5 presents the performance metrics of various classification algorithms applied to the UCI dataset using stratified sampling. The RF algorithm achieved perfect scores across all metrics, with 100% accuracy and zero error rates. DT closely followed with 98.54% accuracy and an error rate of 1.46%. GB performed impressively with 97.56% accuracy. SVM showed 92.68% accuracy, while KNN and GNB achieved accuracies of 86.34% and 82.93%, respectively. LR exhibited the lowest accuracy among the models at 80.98%. Overall, RF demonstrated superior performance, while LR lagged behind.

Moreover, Fig. 5 shows a comparison of error rates for various ML algorithms applied to the UCI dataset using stratified sampling. RF and DT achieve the lowest error rates, indicating the best performance among the evaluated methods. GB also exhibits a relatively low error rate, demonstrating strong predictive capabilities. SVM and KNN fall in the mid-range of error rates. LR and GNB show higher error rates, reflecting comparatively weaker performance. This comparison underscores the effectiveness of RF, DT, and GB in minimizing prediction errors.

Table 6 presents the performance metrics of various classification algorithms applied to the UCI dataset using k-fold cross-validation. LR achieved the highest accuracy of 84.16%, along with the best precision (0.8472) and F1-score (0.8392). RF followed with an accuracy of 81.52%, while SVM slightly outperformed RF in terms of precision (0.8292) and F1-score (0.8188), despite a marginally lower accuracy of 82.17%. GNB and GB exhibited similar performances, both with around 80.84% and 80.87% accuracy, respectively. KNN had a lower accuracy of 79.85% and the highest log loss (1.751), indicating weaker performance compared to other models. DT demonstrated the lowest accuracy (72.89%)

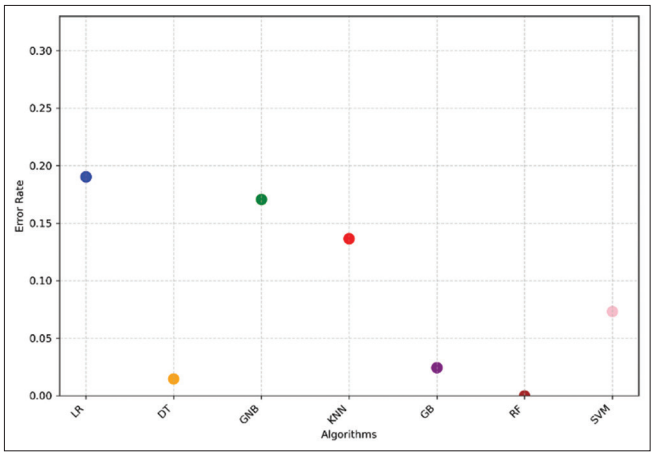


Fig. 5. Comparison of error rates for machine learning algorithms using stratified sampling on UCI dataset.

and the highest log loss (9.773), making it the least effective model in this comparison. Overall, LR emerged as the most robust classifier, balancing accuracy, precision, and F1-score, while DT exhibited the weakest performance due to its high error rate and log loss.

Fig. 6 illustrates the error rates of various ML algorithms applied to the UCI dataset using k-fold cross-validation. The algorithms compared include LR, DT, GNB, KNN, GB, RF, and SVM. The error rate is plotted on the y-axis, while the different algorithms are labeled along the x-axis. For clarity, each algorithm is represented by a distinct color. The DT exhibits the highest error rate among the models, while LR achieves the lowest. The remaining models demonstrate relatively similar error rates, with slight variations. These results provide insights into the comparative performance of different classifiers on the given dataset, helping to determine the most effective model for classification tasks.

The performance metrics of classification algorithms on the HD dataset are shown in Table 7. KNN achieved the highest accuracy (90.16%) and the lowest error rate (9.83%), making it the top-performing algorithm in this study. SVM

TABLE 6: Performance metrics of classification algorithms on the UCI dataset using k-fold cross-validation

Algorithm	Accuracy	Precision	F1-score	Log loss	Error rate
LR	84.15847	0.847168	0.839233	0.408436	0.158415
RF	81.51913	0.819266	0.81326	0.404027	0.184809
SVM	82.17486	0.829186	0.818829	0.420519	0.178251
GNB	80.84153	0.8137	0.806308	0.580949	0.191585
GB	80.86885	0.811012	0.806967	0.476965	0.191311
KNN	79.84699	0.804062	0.794404	1.751016	0.20153
DT	72.88525	0.72869	0.728066	9.773148	0.271148

LR: Logistic regression, RF: Random forest, SVM: Support vector machine, GNB: Gaussian Naive Bayes, GB: Gradient boosting, KNN: K-nearest neighbors, DT: Decision tree

TABLE 7: Performance metrics of classification algorithms on the HD dataset

Algorithm	Accuracy	Precision	F1-score	Log loss	Error rate
LR	85.2459	0.853076	0.852538	0.364798	0.147541
RF	83.60656	0.836066	0.836066	0.3633	0.163934
SVM	86.88525	0.870862	0.868923	0.345582	0.131148
GNB	86.88525	0.870862	0.868923	0.693538	0.131148
GB	78.68852	0.787537	0.787	0.43882	0.213115
KNN	90.16393	0.903684	0.901692	1.96402	0.098361
DT	75.40984	0.770801	0.75224	8.863193	0.245902

LR: Logistic regression, RF: Random forest, SVM: Support vector machine, GNB: Gaussian Naive Bayes, GB: Gradient boosting, KNN: K-nearest neighbors, DT: Decision tree

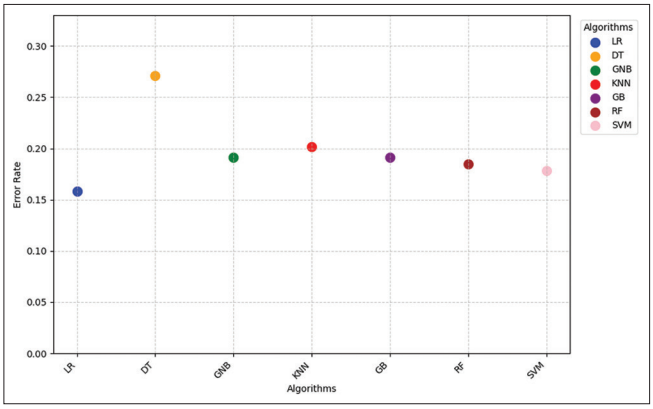


Fig. 6. Comparison of error rates across algorithms for UCI dataset using k-fold cross-validation.

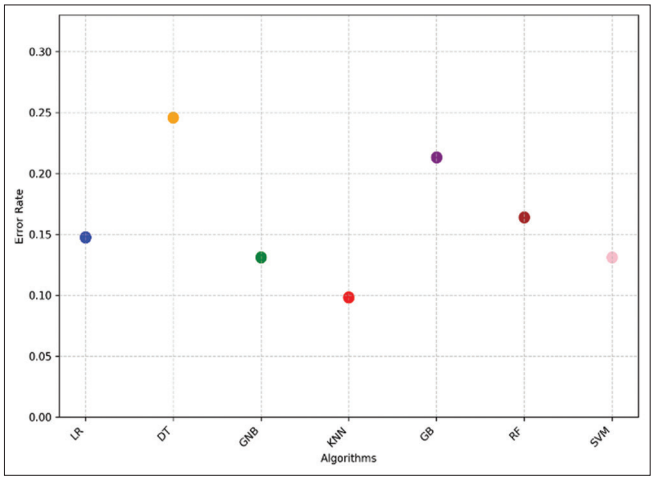


Fig. 7. Comparison of error rates across algorithms for HD dataset classification.

and GNB both exhibited strong performance with identical accuracy (86.89%), precision (0.87), and F1-score (0.87). LR also performed well with an accuracy of 85.25%. RF showed moderate results with an accuracy of 83.61%, while DT achieved comparatively lower accuracy at 78.69% and 75.41%, respectively. The log loss values for SVM and LR were notably lower, indicating better calibration, while DT had the highest log loss, suggesting poorer reliability. These results position KNN as the most effective classifier for this dataset, with SVM and GNB following closely behind.

The classification results demonstrate significant performance variations across different algorithms. KNN achieved the lowest error rate, highlighting its effectiveness for this dataset. LR also performed well, with relatively low error rates. In contrast, DT exhibited the highest error rate, reflecting its lower suitability for this task. RF and GB delivered intermediate performance, balancing error rates and model complexity. SVM showed moderate performance, while GNB lagged behind the top-performing algorithms but outperformed DT. These findings, as depicted in Fig. 7, underscore the importance of selecting an appropriate algorithm to enhance classification accuracy.

The performance metrics of classification algorithms on the HD dataset using stratified sampling are presented in Table 8,

TABLE 8: Performance metrics of classification algorithms on the HD dataset using stratified sampling

Algorithm	Accuracy	Precision	F1-score	Log loss	Error rate
LR	80.32787	0.812564	0.799672	0.438066	0.196721
RF	83.60656	0.858297	0.831266	0.404595	0.163934
SVM	81.96721	0.834563	0.815437	0.425831	0.180328
GNB	81.96721	0.826237	0.817182	0.61126	0.180328
GB	81.96721	0.826237	0.817182	0.449421	0.180328
KNN	80.32787	0.812564	0.799672	2.069169	0.196721
DT	70.4918	0.705287	0.702	10.63583	0.295082

LR: Logistic regression, RF: Random forest, SVM: Support vector machine, GNB: Gaussian Naive Bayes, GB: Gradient boosting, KNN: K-nearest neighbors, DT: Decision tree

highlighting key insights. LR achieved an accuracy of 80.33%, a precision of 0.813, and an F1-score of 0.800, with a log loss of 0.438 and an error rate of 0.197. RF outperformed others with the highest accuracy of 83.61%, precision of 0.858, and an F1-score of 0.831, while maintaining a log loss of 0.405 and an error rate of 0.164. SVM and GB both reached an accuracy of 81.97%, with comparable precision and F1-Scores, though GB had slightly better log loss at 0.449. KNN and GNB had identical accuracies of 80.33% and 81.97%, respectively, but KNN showed significantly higher log loss at 2.069. DT performed the worst, with an accuracy of 70.49% and a log loss of 10.636, reflecting its limitations compared to other models.

Fig. 8 compares the error rates of various classification algorithms for the HD dataset using stratified sampling. DT exhibited the highest error rate, approximately 0.30, highlighting its relatively poor performance. LR, KNN, and GNB demonstrated similar error rates around 0.20, indicating moderate performance. GB, RF, and SVM achieved lower error rates, with RF standing out as the most accurate model, achieving an error rate of approximately 0.16. This emphasizes the effectiveness of ensemble methods, such as RF and GB in reducing classification errors compared to simpler models such as DT.

Table 9 presents the performance metrics of various classification algorithms evaluated on the HD dataset using k-fold cross-validation. The DT algorithm achieved the highest performance, attaining 100% accuracy, precision, and F1-score, with a log loss of 0 and an error rate of 0. RF closely followed, exhibiting an accuracy of 99.61%, precision of 0.9962, and an F1-score of 0.9961, with minimal log loss (0.0516) and a very low error rate (0.0039). GB also demonstrated strong performance, achieving 97.17% accuracy, a precision of 0.9722, and an F1-score of 0.9717, with a log loss of 0.1364 and an error rate of 0.0283. The SVM classifier achieved an accuracy of 92.39% and a precision of 0.9241, with an F1-score of 0.9239 and

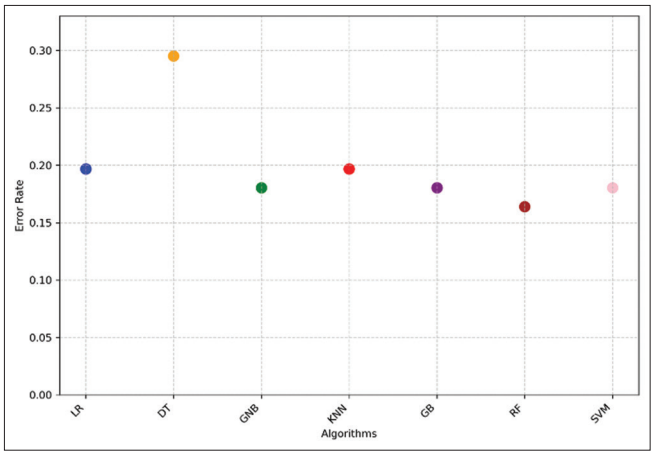


Fig. 8. Comparison of error rates across algorithms for HD dataset classification using stratified sampling.

a moderate log loss of 0.1991. LR and KNN exhibited comparable performance, with 84.59% and 83.61% accuracy, respectively. LR showed a slightly higher precision (0.8506) and F1-score (0.8451) compared to KNN (0.8380 precision, 0.8359 F1-score). GNB recorded the lowest accuracy (82.63%) and the highest log loss (0.5122), indicating weaker reliability than other models. Overall, DT and RF emerged as the most effective classifiers, with GB also demonstrating competitive performance.

Fig. 9 presents a comparative analysis of the error rates for various ML algorithms applied to the HD dataset using k-fold cross-validation. The x-axis represents different algorithms, including LR, DT, GNB, KNN, GB, RF, and SVM. The y-axis denotes the corresponding error rates. For clarity, each algorithm is represented by a distinct color. The results indicate that DT exhibits the lowest error rate, demonstrating its strong classification performance for this dataset. In addition, RF and GB also show relatively low error rates, reinforcing the effectiveness of ensemble methods. Conversely, GNB and KNN yield higher error rates, while LR and SVM fall in between. These findings highlight the

TABLE 9: Performance metrics of classification algorithms on the HD dataset using k-fold cross-validation

Algorithm	Accuracy	Precision	F1-score	Log loss	Error rate
LR	84.58537	0.850558	0.845109	0.362773	0.154146
RF	99.60976	0.996248	0.996098	0.051581	0.003902
SVM	92.39024	0.924054	0.923882	0.199138	0.076098
GNB	82.63415	0.828886	0.825682	0.512226	0.173659
GB	97.17073	0.97216	0.971708	0.136384	0.028293
KNN	83.60976	0.837965	0.835948	0.222265	0.163902
DT	100	1	1	0	0

LR: Logistic regression, RF: Random forest, SVM: Support vector machine, GNB: Gaussian Naive Bayes, GB: Gradient boosting, KNN: K-nearest neighbors, DT: Decision tree

TABLE 10: Performance metrics of classification algorithms on the UCI-HD dataset

Algorithm	Accuracy	Precision	F1-score	Log loss	Error rate
LR	83.08271	0.838173	0.829144	0.377472	0.169173
RF	100.00000	1.000000	1.000000	0.025382	0.000000
SVM	92.85714	0.930338	0.928379	0.183863	0.071429
GNB	82.70677	0.832169	0.825731	0.543392	0.172932
GB	96.99248	0.970311	0.969898	0.128287	0.030075
KNN	94.73684	0.94913	0.947392	0.129948	0.052632
DT	100.0000	1.0000	1.0000	2.22E-16	0.000000

LR: Logistic regression, RF: Random forest, SVM: Support vector machine, GNB: Gaussian Naive Bayes, GB: Gradient boosting, KNN: K-nearest neighbors, DT: Decision tree

performance variations among different models and suggest that DT is particularly well-suited for this dataset.

Significantly, this model was modified to show that the combination of these two datasets affected its performance. Therefore, Table 10 presents the performance metrics of various classification algorithms on the UCI-HD dataset, offering insights into their effectiveness. DT and RF both achieved perfect performance with 100% accuracy, precision, and F1-score, along with zero error rates and minimal log loss. GB exhibited excellent results, with 96.99% accuracy and a precision of 0.970, followed by KNN at 94.73% accuracy. SVM showed strong performance, achieving 92.86% accuracy and a log loss of 0.1838. LR and GNB demonstrated moderate effectiveness, with accuracies of 83.08% and 82.71%, respectively. The table highlights the dominance of ensemble methods and DT in classification tasks.

The error rates of various ML algorithms applied to the UCI Heart Disease (UCI-HD) dataset are illustrated in Fig. 10. RF and DT achieved the lowest error rates, showcasing their strong predictive performance. GB also performed competitively, with slightly higher error rates. KNN algorithm demonstrated moderate accuracy, while LR and GNB produced comparatively higher error rates, indicating reduced effectiveness.

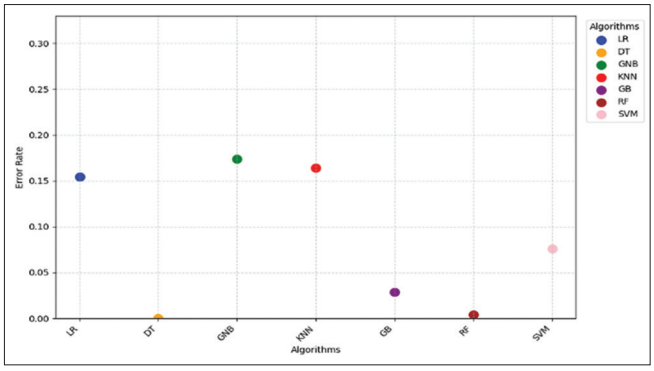


Fig. 9. Comparison of error rates across algorithms for the HD dataset using k-fold cross-validation.

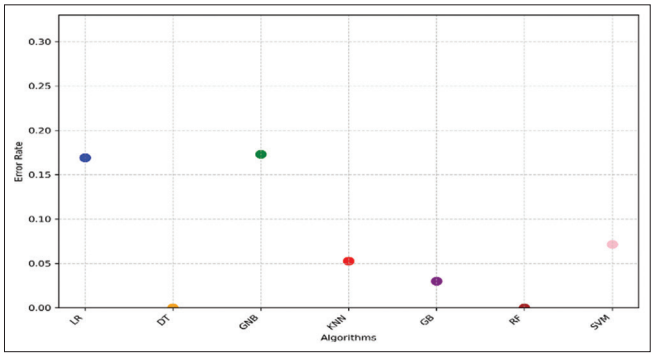


Fig. 10. Comparison of error rates for machine learning algorithms on the UCI-HD dataset.

The results clearly determine the performance metrics of various classification algorithms applied to the UCI-HD dataset using stratified sampling, as shown in Table 11. DT and RF achieved perfect results, with 100% accuracy,

TABLE 11: Performance metrics of classification algorithms on UCI-HD dataset using stratified sampling

Algorithm	Accuracy	Precision	F1-score	Log loss	Error rate
LR	84.58647	0.847779	0.845358	0.352075	0.154135
RF	100.00000	1.000000	1.000000	0.021479	0.000000
SVM	95.11278	0.95114	0.95112	0.140181	0.048872
GNB	84.21053	0.842237	0.84198	0.47266	0.157895
GB	98.1203	0.981225	0.9812	0.12264	0.018797
KNN	96.61654	0.96678	0.966129	0.116318	0.033835
DT	100.00000	1.000000	1.000000	2.22E-16	0.000000

LR: Logistic regression, RF: Random forest, SVM: Support vector machine, GNB: Gaussian Naive Bayes, GB: Gradient boosting, KNN: K-nearest neighbors, DT: Decision tree

precision, and F1-score, and zero error rates. GB and KNN also demonstrated high performance, with accuracies of 98.12% and 96.62%, respectively. SVM showed strong results, achieving 95.11% accuracy. LR and GNB exhibited moderate performance, with accuracies of 84.59% and 84.21%. The metrics highlight the superior predictive capabilities of ensemble models, such as RF and DT.

The results indicate that for the stratified modification, the error rates of various classification algorithms applied to the UCI-HD dataset using stratified sampling are illustrated in Fig. 11. The comparison reveals that SVM achieved the lowest error rate, indicating its superior performance in accurately classifying the dataset. RF and GB also exhibited competitive error rates, showcasing their effectiveness in handling the dataset's complexity. Conversely, algorithms such as GNB demonstrated relatively higher error rates, suggesting challenges in capturing the underlying data patterns. The stratified sampling methodology ensured balanced class representation, which contributed to the robustness of the evaluation. These results emphasize the need for selecting appropriate algorithms for high-stakes applications, such as heart disease prediction, where classification accuracy is paramount.

Table 12 presents the performance metrics of various classification algorithms evaluated on the UCI-HD dataset using k-fold cross-validation. The DT and RF classifiers achieved perfect accuracy (100%) with an error rate of 0, demonstrating their strong predictive capabilities. The GB classifier followed closely with an accuracy of 98.12%, a precision of 0.981, and an F1-score of 0.981, indicating robust performance. The SVM model also exhibited high accuracy (94.88%) and precision (0.949), while KNN performed slightly lower with an accuracy of 93.53%. LR and GNB recorded relatively lower accuracy scores of 85.16% and 82.91%, respectively, with GNB having the highest log loss value (0.5147), indicating more significant uncertainty in its predictions. DT achieved the lowest log loss among all

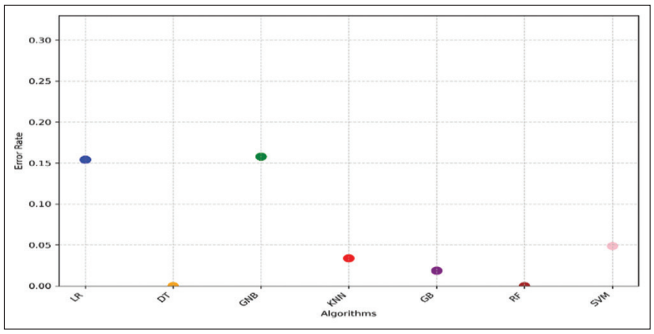


Fig. 11. Comparison of error rates across algorithms for UCI-HD dataset classification using stratified sampling.

models, emphasizing its reliability. Overall, ensemble-based methods (RF and GB) outperformed other classifiers in terms of accuracy and precision, highlighting their effectiveness in heart disease classification.

Fig. 12 presents a comparative analysis of the error rates for different ML algorithms on the UCI-HD dataset using k-fold cross-validation. The x-axis represents the algorithms evaluated, including LR, DT, GNB, KNN, GB, RF, and SVM. The y-axis denotes the corresponding error rates. Each algorithm is color-coded for clarity, as shown in the legend. The results indicate that RF and GB achieved the lowest error rates, suggesting superior predictive performance, while GNB and LR exhibited higher error rates. The variability in error rates highlights the importance of model selection in achieving optimal classification performance on this dataset.

5. DISCUSSION

There are differences in the evaluation models and results obtained from the combined datasets or stratified algorithms compared to previous studies on the same datasets. All relevant studies are referenced in the background review. However, the variations in performance highlight the superior results achieved by specific algorithms, as noted.

TABLE 12: Performance metrics of classification algorithms on the UCI-HD dataset using k-fold cross-validation

Algorithm	Accuracy	Precision	F1-score	Log loss	Error rate
LR	85.16385303	0.857201168	0.850471332	0.361969257	0.14836147
RF	100	1	1	0.020635519	0
SVM	94.87955738	0.949310038	0.948753901	0.15644934	0.051204426
GNB	82.90707902	0.830871801	0.828611865	0.514689846	0.17092921
GB	98.11717974	0.981378517	0.981166354	0.123583547	0.018828203
KNN	93.52617393	0.936109038	0.935228791	0.1410065	0.064738261
DT	100	1	1	0	0

LR: Logistic regression, RF: Random forest, SVM: Support vector machine, GNB: Gaussian Naive Bayes, GB: Gradient boosting, KNN: K-nearest neighbors, DT: Decision tree

Table 13 presents the accuracy of several classification algorithms across the UCI, HD, and combined UCI-HD datasets. LR achieved accuracies of 79.512%, 85.246%, and 83.083% for the UCI, HD, and UCI-HD datasets, respectively. RF demonstrated the highest accuracy on the UCI dataset (98.537%) but showed a decline on the HD dataset (83.607%), achieving a perfect accuracy of 100% on the combined dataset. SVM recorded 88.78% for UCI, 83.607% for HD, and 92.857% for UCI-HD. GNB achieved 80% on UCI, 86.885% on HD, and 82.707% on the combined dataset. GB performed with 93.171% on UCI, 78.689% on HD, and 96.992% on UCI-HD. KNN achieved 83.415% on UCI, 90.164% on HD, and 94.737% on the combined dataset. Finally, DT reached an accuracy of 98.537% on UCI, 75.41% on HD, and 100% on UCI-HD. Overall, RF and DT demonstrated the best performance on the combined UCI-HD dataset, achieving perfect accuracy. However, other algorithms, such as GB and KNN, also showed notable effectiveness across the datasets.

TABLE 13: Accuracy of classification algorithms on UCI, HD, and combined UCI-HD datasets

Algorithm	Accuracy UCI	Accuracy HD	Accuracy UCI-HD
LR	79.512	85.246	83.083
RF	98.537	83.607	100
SVM	88.78	83.607	92.857
GNB	80	86.885	82.707
GB	93.171	78.689	96.992
KNN	83.415	90.164	94.737
DT	98.537	75.41	100

LR: Logistic regression, RF: Random forest, SVM: Support vector machine, GNB: Gaussian Naive Bayes, GB: Gradient boosting, KNN: K-nearest neighbors, DT: Decision tree

The comparison in Table 14 highlights the impact of using stratification on algorithm accuracy with the UCI dataset. Without stratification, the accuracy varied across algorithms, with DT and RF achieving the highest accuracies at 98.537%, reflecting their strong classification capabilities. GB also performed well at 93.171%, while SVM, KNN, and GNB exhibited moderate accuracies of 88.780%, 83.415%, and 80%, respectively. LR showed the lowest accuracy at 79.512%, indicating potential limitations with non-stratified data. Stratification improved accuracy for most algorithms, ensuring better representation of class distributions during training. RF achieved a perfect 100% accuracy with stratification, while GB, SVM, and KNN also demonstrated significant gains. DT performance remained constant, indicating minimal dependency on class distribution in this case.

TABLE 14: Comparison of algorithm accuracies with and without stratification according to the UCI dataset

Algorithm	Accuracy	Accuracy (Stratify=y)
LR	79.512	80.976
RF	98.537	100.000
SVM	88.780	92.683
GNB	80.000	82.927
GB	93.171	97.560
KNN	83.415	86.341
DT	98.537	98.537

LR: Logistic regression, RF: Random forest, SVM: Support vector machine, GNB: Gaussian Naive Bayes, GB: Gradient boosting, KNN: K-nearest neighbors, DT: Decision tree

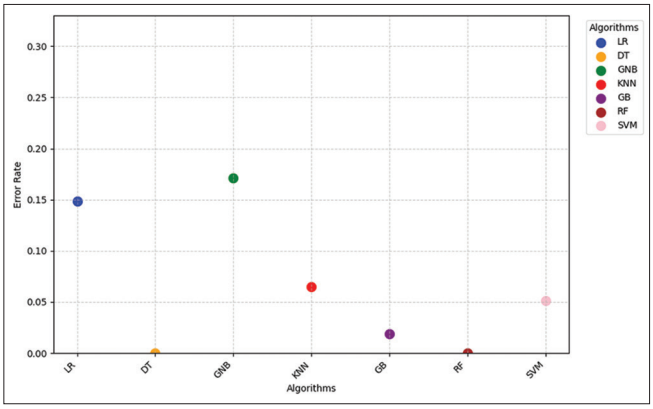


Fig. 12. Comparison of error rates across algorithms for the UCI-HD dataset using k-fold cross-validation.

Table 15 compares algorithm accuracies on the HD dataset with and without stratified sampling. Most algorithms

showed a decline in accuracy with stratified sampling, such as KNN (90.164–80.328%) and LR (85.246–80.328%), indicating sensitivity to data redistribution. RF maintained consistent accuracy (83.607%) across both cases, showcasing its robustness. GB slightly improved (78.689–81.967%), suggesting enhanced generalization. Other algorithms, such as SVM and GNB, experienced moderate declines, indicating varied sensitivity. These results highlight that stratified sampling impacts algorithms differently, emphasizing the need for careful evaluation of sampling strategies for optimal performance.

The results in Table 16 highlight the impact of stratified sampling on algorithm accuracy when applied to the UCI-HD dataset. The majority of algorithms show an improvement in accuracy with stratified sampling, notably LR, which increases from 83.083% to 84.586%, and GNB, which improves from 82.707% to 84.211%. Similarly, KNN shows a notable increase from 94.737% to 96.617%, and GB improves from 96.992% to 98.120%. SVM also sees an increase from 92.857% to 95.113%. In contrast, DT and RF algorithms maintain their perfect accuracy of 100% with and without stratification. These findings suggest that stratified sampling can provide slight accuracy enhancements, particularly for algorithms that initially perform below perfect accuracy, while having no effect on algorithms already achieving optimal results. This suggests the potential of stratified sampling to improve model performance, especially for classifiers sensitive to class imbalances.

TABLE 15: Algorithm accuracies: Comparison with and without stratified sampling HD dataset

Algorithm	Accuracy	Accuracy (Stratify=y)
LR	85.246	80.328
RF	83.607	83.607
SVM	83.607	83.607
GNB	86.885	81.967
GB	78.689	81.967
KNN	90.164	80.328
DT	75.410	70.492

LR: Logistic regression, RF: Random forest, SVM: Support vector machine, GNB: Gaussian Naive Bayes, GB: Gradient boosting, KNN: K-nearest neighbors, DT: Decision tree

TABLE 16: Algorithm accuracies: Comparison with and without stratified sampling UCI – HD dataset

Algorithm	Accuracy	Accuracy (Stratify=y)
LR	83.083	84.586
RF	100	100
SVM	92.857	95.113
GNB	82.707	84.211
GB	96.992	98.120
KNN	94.737	96.617
DT	100	100

LR: Logistic regression, RF: Random forest, SVM: Support vector machine, GNB: Gaussian Naive Bayes, GB: Gradient boosting, KNN: K-nearest neighbors, DT: Decision tree

In previous experiments with the UCI and HD datasets, several machines learning algorithms, including LR, RF, SVM, GNB, and KNN, were evaluated. For the UCI dataset, RF achieved the highest accuracy of 98.53%, followed by SVM at 87.31%. LR reached an accuracy of 82.92%, while KNN recorded 81.95%. GNB showed the lowest accuracy at 74.63% [34]. Notably, our modifications resulted in improved accuracy across these models compared to previous results.

Table 17 presents the classification accuracy of seven algorithms, LR, RF, SVM, GNB, GB, KNN, and DT, evaluated on the UCI, HD, and combined UCI-HD datasets using stratified sampling. Among these, RF and DT achieved perfect accuracy (100%) on both the UCI and UCI-HD datasets, indicating excellent performance. SVM also delivered strong results, with accuracy scores of 92.68% on UCI and 95.11% on UCI-HD, closely followed by GB with 97.56% and 98.12%, respectively.

However, DT showed a sharp decline in performance on the HD dataset, recording the lowest accuracy of 70.49%, while RF maintained a comparatively higher score of 83.61%. This suggests that the HD dataset introduced more complex classification challenges for tree-based models. Overall, the combined UCI-HD dataset resulted in improved accuracy for most algorithms, demonstrating the advantage of data integration. LR, GNB, and KNN exhibited moderate yet stable performance across all datasets, with accuracies ranging from 80.33% to 96.62%. These findings underscore the superior generalization ability of ensemble methods such as RF and GB, particularly when applied to enriched and well-balanced datasets.

The exceptionally high accuracy (100%) observed for the RF and DT models, particularly on the combined UCI-HD dataset with stratified sampling, may initially raise concerns of overfitting. However, several factors inherent to the study provide a reasonable justification for this performance. First, the application of stratified sampling ensured balanced class representation across both training and testing sets, thereby minimizing the risk of class imbalance-related bias. Second, the datasets underwent rigorous pre-processing, which eliminated missing values and ensured well-defined feature distributions, likely enhancing the learning capacity of tree-based algorithms. Third, the combined dataset incorporated consistent features from two closely related sources

TABLE 17: Accuracy of classification algorithms on UCI, HD, and combined UCI-HD datasets using stratified sampling

Algorithm	Accuracy UCI (Stratify=y)	Accuracy HD (Stratify=y)	Accuracy UCI-HD (Stratify=y)
LR	80.976	80.328	84.586
RF	100.000	83.607	100
SVM	92.683	83.607	95.113
GNB	82.927	81.967	84.211
GB	97.560	81.967	98.120
KNN	86.341	80.328	96.617
DT	98.537	70.492	100

LR: Logistic regression, RF: Random forest, SVM: Support vector machine, GNB: Gaussian Naive Bayes, GB: Gradient boosting, KNN: K-nearest neighbors, DT: Decision tree

(UCI and HD), which may have facilitated more distinct classification boundaries. Moreover, the RF, as an ensemble method, mitigates overfitting by averaging the outputs of multiple decorrelated DTs. While perfect accuracy warrants cautious interpretation, the results remained consistent across various datasets and sampling techniques, suggesting that the model's performance is robust rather than indicative of data memorization. Nonetheless, to further confirm generalizability, additional validation on external datasets or through k-fold cross-validation would be a valuable next step in future work. It is also possible that data augmentation was employed to prevent overfitting.

The study evaluated ML models for CVD prediction, showing distinct performance patterns. Ensemble models, such as RF, GB, and DT were top performers, with RF achieving perfect accuracy, especially using stratified sampling. Stratification was crucial for improving SVM and GNB by maintaining class balance and reducing bias. LR showed moderate performance due to its linear limitations, while SVM excelled in high-dimensional spaces. KNN performed well on smaller datasets but struggled with larger, complex ones. GNB was competitive but limited by its Gaussian assumptions. Combining UCI and HD datasets enhanced performance for all models, especially ensemble methods. RF and DT achieved 100% accuracy on the combined dataset. Despite perfect accuracy, safeguards, such as pre-processing and stratification minimized overfitting risks. The findings highlight the value of ensemble methods, dataset integration, and stratified sampling, suggesting future work on hybrid models and real-world validation.

Fig. 13 displays the classification accuracy of seven ML algorithms, LR, RF, SVM, GNB, GB, KNN, and DT, evaluated using k-fold cross-validation. Performance is compared across three datasets: UCI, HD, and their merged variant (UCI-HD). Notably, RF and DT achieved perfect accuracy (100%) on the HD and UCI-HD datasets, while

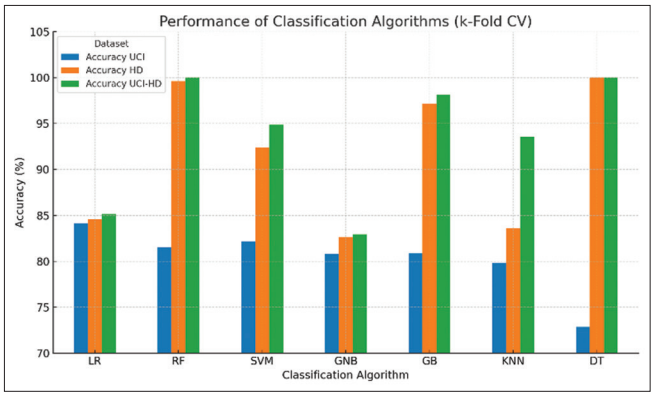


Fig. 13. Comparison of classification algorithm performance across UCI, HD, and UCI-HD datasets using k-Fold cross-validation.

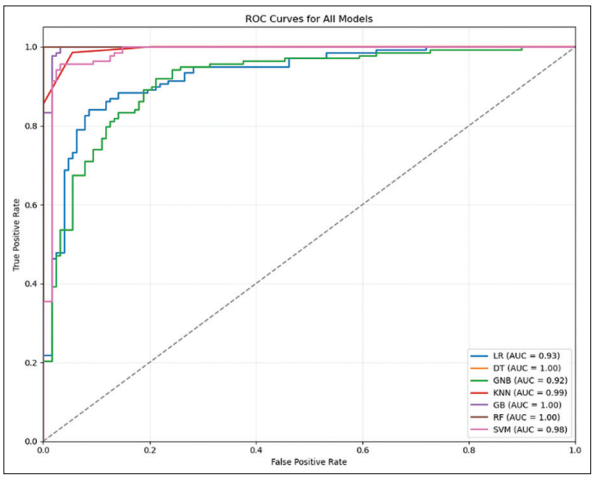


Fig. 14. ROC curve evaluation of machine learning models: Ensemble methods achieve perfect AUC.

LR and SVM demonstrated consistently high performance across all datasets. The figure highlights the impact of dataset variation on model accuracy, underscoring the robustness of ensemble methods, such as RF and GB.

Fig. 14 illustrates the Receiver Operating Characteristic (ROC) curves for all evaluated classification models, providing a

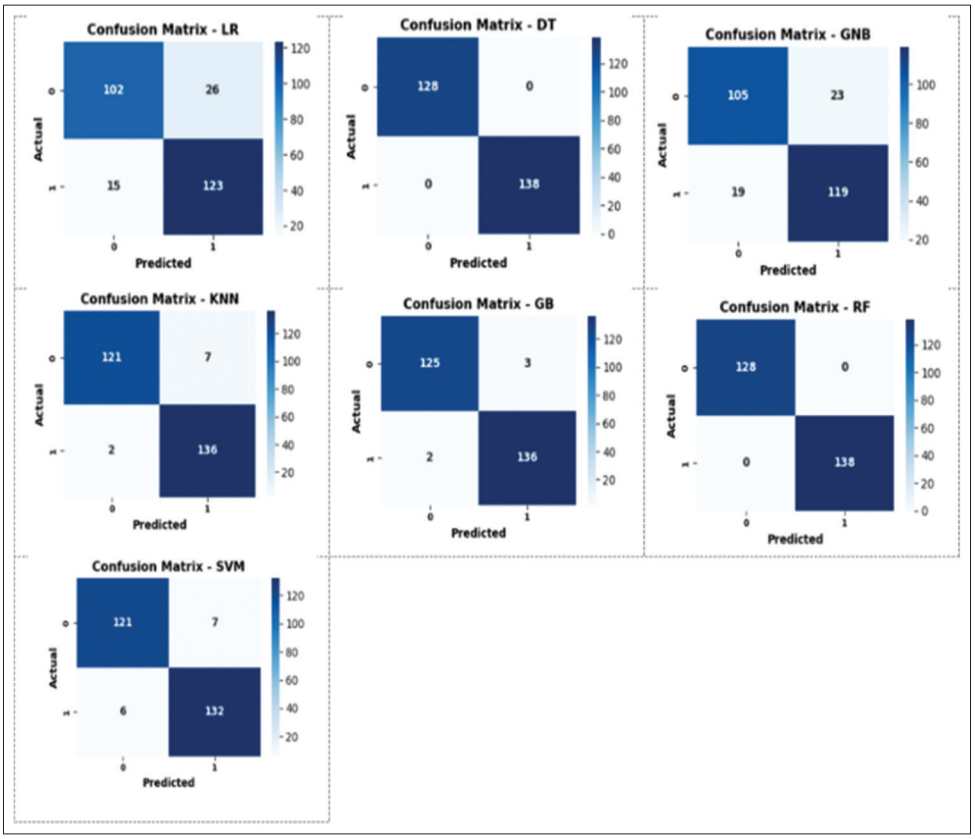


Fig. 15. Confusion matrix comparison of ML models.

visual comparison of their diagnostic performance. The Area Under the Curve (AUC) metric is presented for each model, with higher values indicating better discriminative ability. Ensemble methods, such as DT, GB, and RF achieved perfect classification performance with an AUC of 1.00. The KNN and SVM models also demonstrated excellent performance with AUCs of 0.99 and 0.98, respectively. LR and GNB had slightly lower AUCs of 0.93 and 0.92, but still showed strong classification capability. Overall, the ROC analysis confirms the superior performance of ensemble methods and supports their robustness in distinguishing between classes with minimal false positive rates.

The comparative analysis of confusion matrices, as illustrated in Fig. 15, reveals substantial differences in classification performance across the evaluated models. Both DT and RF classifiers achieved perfect predictive accuracy (100%) with F1-scores of 1.00, indicating zero misclassifications for both negative (class 0) and positive (class 1) instances, 128 and 138 samples, respectively. GB also demonstrated exceptional performance, attaining an accuracy of 98.9% and an F1-score of 0.989, with only five misclassifications (three FP and two FN). Similarly, the KNN classifier performed strongly, with 97.8% accuracy and a 0.978 F1-score, misclassifying seven negative and two positive cases.

The SVM model yielded commendable results with 96.9% accuracy and a 0.969 F1-score, though it incurred slightly higher misclassification rates (seven FP and six FN). In contrast, LR and GNB underperformed relative to the other models. LR achieved 89.9% accuracy and a 0.896 F1-score, misclassifying 26 negative and 15 positive instances. GNB recorded the lowest performance, with an accuracy of 87.0% and an F1-score of 0.873, resulting in 42 total misclassifications.

These results, as depicted in Fig. 15, underscore the superiority of ensemble methods, particularly RF and GB, in effectively capturing complex patterns within the dataset. Conversely, simpler linear (LR) and probabilistic GNB models may be less capable in such high-dimensional classification tasks.

6. CONCLUSION

CVD remains a leading cause of mortality worldwide, emphasizing the critical need for early diagnosis and

intervention. If the symptoms of heart disease are not promptly identified and treated, the condition can escalate into life-threatening scenarios. Artificial intelligence has been effectively utilized for CVD prediction, with advancements promising increasingly accurate forecasts based on historical medical data. Despite significant progress in this domain, continuous enhancements in predictive methodologies are essential and highly encouraged.

In conclusion, these modifications improved CVD prediction by utilizing seven ML algorithms: LR, RF, SVM, GNB, GB, KNN, and DT. By analyzing medical histories of patients with severe heart conditions, the models classified individuals based on their risk of developing CVD. The models were trained and tested on datasets containing factors such as chest discomfort, high blood pressure, and cardiac arrest. To ensure robustness, evaluations were performed with and without the stratify parameter. The results revealed that the DT and RF algorithms consistently achieved the peak accuracy rates, with both models reaching 100% accuracy on the combined dataset. Besides, the stratified technique enhanced the accuracy across all methods. These findings emphasize the critical role of sufficient training data and stratification in improving predictive performance. They also highlight the potential of AI-driven tools to assist healthcare professionals in making faster and more accurate diagnoses, ultimately lowering costs and enhancing patient outcomes. The results represent a significant advancement in the field by achieving higher accuracy rates than previous studies, setting a standard for the practical application of ML in CVD prediction. Among the tested methods, DT and RF emerged as the most reliable, highlighting the efficiency of ensemble learning techniques in clinical applications.

This research inspires future work to explore integrating additional datasets, refining algorithms, and developing real-time prediction systems to further advance the field. In addition, the combined dataset (UCI-HD) was tested using novel classifier metaheuristic algorithms, such as the fitness dependent optimizer (FDO) [35], [36].

REFERENCES

[1] X. Han. "Heart Disease Type Prediction Model Based on SVM-ANN". In: *Proceedings of the 2022 6th International Conference on Electronic Information Technology and Computer Engineering*, pp. 422-426, 2022.

[2] A. R. Snigdha, S. N. Tasnim, K. R. Miah and T. Islam. "Early Prediction of Heart Attack using Machine Learning Algorithms". In: *Proceedings of the 2nd International Conference on Computing Advancements*, pp. 344-348, 2022.

[3] A. Lahsasna, R. N. Ainon, R. Zainuddin and A. Bulgiba. "Design of a fuzzy-based decision support system for coronary heart disease diagnosis". *Journal of Medical Systems*, vol. 36, pp. 3293-3306, 2012.

[4] S. Song, T. Chen and G. Antoniou. "ANFIS Models for Heart Disease Prediction". In: *Proceedings of the 2021 5th International Conference on Innovation in Artificial Intelligence*", pp. 32-35, 2021.

[5] T. Suresh, T. A. Assegie, S. Rajkumar and N. Komal Kumar. "A hybrid approach to medical decision-making: Diagnosis of heart disease with machine-learning model". *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 12, no. 2, p. 1831, 2022.

[6] A. A. Hussein. "Improve the performance of K-means by using genetic algorithm for classification heart attack". *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 8, no. 2, p. 1256, 2018.

[7] K. Wang, J. Tian, C. Zheng, H. Yang, J. Ren, Y. Liu and Q. Han, Y. Zhang. "Interpretable prediction of 3-year all-cause mortality in patients with heart failure caused by coronary heart disease based on machine learning and SHAP". *Computers in Biology and Medicine*, vol. 137, p. 104813, 2021.

[8] S. Geetha, C. P. Devi, V. Kalaivani, C. J. Haritha and G. Preetha. "Prediction techniques of heart disease and diabetes disease using machine learning". *Turkish Journal of Computer and Mathematics Education*, vol. 12, no. 10, pp. 3316-3325, 2021.

[9] D. O. Hasan and A. M. Aladdin. "Sleep-related consequences of the COVID-19 pandemic: A survey study on insomnia and sleep apnea among affected individuals". *Insights in Public Health Journal*, vol. 5, no 2, 2024.

[10] R. K. Muhammed, R. R. Aziz, A. A. Hassan, A. M. Aladdin, S. J. Saydahet and T. A. Rashidal. "Comparative analysis of AES, blowfish, twofish, salsa 20, and ChaCha20 for image encryption". *Kurdistan Journal of Applied Research*, vol. 9, no. 1, pp. 52-65, 2024.

[11] Z. Rayan, M. Alfonse and A. B. M. Salem. "Machine learning approaches in smart health". *Procedia Computer Science*, vol. 154, pp. 361-368, 2019.

[12] A. M. Aladdin and T. A. Rashid. "*Leo: Lagrange Elementary Optimization*". Germany, Springer, 2024.

[13] A. M. Aladdin and T. A. Rashid. "A new lagrangian problem crossover-a systematic review and meta-analysis of crossover standards". *Systems*, vol. 11, no. 3, p. 144, 2023.

[14] R. Mohammed, N. K. Al-Salihi, T. A. Rashid, A. M. Aladdin, M. Mohammadi and J. Majidpour. "*Artificial Cardiac Conduction System: Simulating Heart Function for Advanced Computational Problem Solving*". [Preprint], 2024.

[15] A. Budianto, R. Ariyuana, and D. Maryono, "Perbandingan K-Nearest Neighbor (Knn) Dan Support Vector Machine (Svm) Dalam Pengenalan Karakter Plat Kendaraan Bermotor," Jurnal Universitas Sebelas Maret, vol. 11, no. 1, p. 27, Nov. 2019, doi: 10.20961/jiptek.v11i1.18018.

[16] A. Gavhane, G. Kokkula, I. Pandya and K. Devadkar. "Prediction of Heart Disease using Machine Learning". In: *2018nd International Conference on Electronics, Communication and Aerospace Technology (ICECA)*, IEEE, 2018, pp. 1275-1278.

[17] S. Ambekar and R. Phalnikar. "Disease Risk Prediction by Using Convolutional Neural Network". In: *2018 4th International Conference on Computing Communication Control and Automation (ICCUBEA)*, IEEE, 2018, pp. 1-5.

[18] N. Jothi, W. Husain, N. A. Rashid and S. Syed-Mohamad.

"Feature selection method using genetic algorithm for medical dataset". *International Journal on Advanced Science Engineering Information Technology*, vol. 9, no. 6, pp. 1907-1912, 2019.

[19] T. A. Assegie. "A support vector machine based heart disease prediction". *Journal of Software Engineering and Intelligent Systems*, vol. 4, pp. 111-116, 2019.

[20] E. S. Kajal and M. Nishika. "Prediction of heart disease using data mining techniques". *International Journal of Advance Research, Ideas and Innovations in Technology*, vol. 2, no. 3, pp. 1-7, 2016.

[21] S. Babu, E. M. Vivek, K. P. Famina, K. Fida, P. Aswathi, M. Shanid and M. Hena. "Heart Disease Diagnosis using Data Mining Technique". In: 2017 *International Conference of Electronics, Communication and Aerospace Technology (ICECA)*. IEEE, 2017, pp. 750-753.

[22] R. Kannan and V. Vasanthi. "Machine Learning Algorithms with ROC Curve for Predicting and Diagnosing the Heart Disease". In: N. B. Muppalaneni, M. Ma and S. Gurumoorthy, Eds. *Soft Computing and Medical Bioinformatics*, Springer, Singapore, 2019, pp. 63-72.

[23] K. Raza. "Improving the Prediction Accuracy of Heart Disease with Ensemble Learning and Majority Voting Rule". In: *U-Healthcare Monitoring Systems*. Academic Press, United States, 2019, pp. 179-196.

[24] L. Sapra, J. K. Sandhu and N. Goyal. "Intelligent method for detection of coronary artery disease with ensemble approach". In: *Advances in Communication and Computational Technology: Select Proceedings of ICACCT 2019*. Springer, 2021, pp. 1033-1042.

[25] A. Al Ahdal, M. Rakhra, R. R. Rajendran, F. Arslan, M. A. Khder, B. Patel and B. R. Rajagopal, R. Jain. "Monitoring cardiovascular problems in heart patients using machine learning". *Journal of Healthcare Engineering*, vol. 2023, no. 1, p. 9738123, 2023.

[26] S. Patidar, A. Jain and A. Gupta. "Comparative Analysis of Machine Learning Algorithms for Heart Disease Predictions". In: *2022 6th International Conference on Intelligent Computing and Control Systems (ICICCS)*, 2022, pp. 1340-1344. doi: 10.1109/ICICCS53718.2022.9788408

[27] N. S. Noori, B. H. Hameed, and M. Kh. Mohammed, "An economic evaluation of the performance efficiency of conservation agriculture and food security projects using logistic regression in iraq for the 2022-2023 season," anbar journal of agricultural sciences, vol. 22, no. 2, pp. 1033–1049, Dec. 2024, doi: 10.32649/ajas.2024.184466.

[28] Y. Chen, L. Li, W. Li, Q. Guo, Z. Du and Z. Xu. "Fundamentals of neural networks". *AI Computing Systems*. Elsevier, Netherlands, pp. 17-51, 2024.

[29] M. Schonlau and R. Y. Zou. "The random forest algorithm for statistical learning". *The Stata Journal: Promoting Communications on Statistics and Stata*, vol. 20, no. 1, pp. 3-29, 2020.

[30] A. U. Haq, J. P. Li, M. H. Memon, S. Nazir and R. Sun. "A hybrid intelligent system framework for the prediction of heart disease using machine learning algorithms". *Mobile Information Systems*, vol. 2018, no. 1, p. 3860146, 2018.

[31] S. Naiem, A. E. Khedr, A. M. Idrees and M. I. Marie. "Enhancing the efficiency of gaussian naïve bayes machine learning classifier in the detection of DDOS in cloud computing". *IEEE Access*, vol. 11, pp. 124597-124608, 2023.

[32] M. Malohlava and A. Candel. "*Gradient Boosting Machine with H2O*". H20 Booklet, 2016. Available from: https://docs.h2o.ai/h2o/latest-stable/h2o-docs/booklets [Last accessed on 2025 Apr 04].

[33] I. Maryani, Rousyati, Indriyanti, D. Pratmanto, Y. M. Kristania and M. Maulidah. "Prediction of Heart Disease using Decision Tree in Comparison with Particle Swarm Optimization to Improve Accuracy". In: *Proceedings of the 3rd International Conference on Advanced Information Scientific Development, SCITEPRESS - Science and Technology Publications*, 2023, pp. 233-239.

[34] S. Patidar, D. Kumar and D. Rukwal. Comparative Analysis of Machine Learning Algorithms for Heart Disease Prediction". In: *ITM Web of Conferences*, 2022. doi: 10.3233/ATDE220723

[35] A. M. Aladdin and A. M. Abdulla. "Fitness-Dependent Optimizer for IoT Healthcare Using Adapted Parameters: A Case Study Implementation". In: *Practical Artificial Intelligence for Internet of Medical Things*, CRC Press, United States, 2023, pp. 45-61.

[36] J. M. Abdullah and T. Ahmed. "*Fitness Dependent Optimizer: Inspired by the Bee Swarming Reproductive Process*". Vol. 7. IEEE Access, Park Avenue, pp. 43473-43486, 2019.

ORIGINAL RESEARCH ARTICLE UHD JOURNAL OF SCIENCE AND TECHNOLOGY

Evaluating the Effectiveness of Traffic Metering Strategies in Reducing Congestion: A Case Study of Amman

Qeethara Al-Shayea, Huthaifa Aljawazneh

Department of Business Intelligence, Faculty of Business, Al-Zaytoonah University of Jordan, Amman, Jordan

ABSTRACT

Traffic congestion is a significant issue in urban road networks, particularly in Amman, where peak hours cause major delays for commuters. Developing an advanced traffic management system is essential to helping residents save time, reduce congestion, and alleviate traffic jams. To address this challenge, we have implemented a simulation model powered by machine learning techniques to effectively and accurately manage traffic flow on Amman's streets. This innovative system leverages real-world data from the Jordanian capital to dynamically optimize traffic control. By automating traffic management processes, the model aims to reduce congestion while easing the workload of traffic personnel. This approach promises to enhance urban mobility and contribute to building a smarter and more efficient traffic management infrastructure in Amman, ensuring a better quality of life for its residents. After implementing the metering strategy, the traffic flow became more balanced, with less congestion and smoother transitions between intersections. The metering points effectively regulated the entry of vehicles into the circles, preventing congestion buildup and improving overall traffic efficiency.

Index Terms: Traffic Congestion, Machine Learning, Simulation, Aimsun Software

1. INTRODUCTION

Traffic congestion in Amman is a complex issue driven by rapid urbanization, population growth, and an aging road network unable to meet the city's evolving transportation demands. The primary cause of this congestion lies in the imbalance between the road network's capacity and the increasing traffic demand. Amman's road infrastructure, characterized by narrow thoroughfares and limited lanes, struggles to accommodate the rising number of vehicles, particularly during peak hours. This challenge is exacerbated by the lack of alternative routes to divert traffic during periods of high congestion [1].

Implementing an intelligent traffic management system is critical to addressing this pressing issue. Such a system would help optimize traffic flow, reduce delays, and alleviate the city's road network strain. The main objective of this study is to reduce the chances of traffic jams forming. This work employs Aimsun Next to simulate and analyze traffic demand under different conditions. The model incorporates real-world traffic data. Aimsun Next is a dedicated microscopic traffic simulation software, which helps in detail modeling of traffic flow alongside vehicle behavior and intersection dynamics for the simulation.

The remainder of this paper is organized as follows: Section 2 reviews the literature review, Section 3 details the methodology, Section 4 submits the results and discussions, and Section 5 argues the findings and conclusions.

Access this article online

DOI: 10.21928/uhdjst.v9n1y2025.pp169-180

E-ISSN: 2521-4217
P-ISSN: 2521-4209

Corresponding author's e-mail: Qeethara Al-Shayea, Business Intelligence Department, Faculty of Business, Al-Zaytoonah University of Jordan, Amman, Jordan. E-mail: drqeethara@zuj.edu.jo

Received: 07-03-2025 Accepted: 10-05-2025 Published: 12-06-2025

2. LITERATURE REVIEW

Numerous studies have investigated strategies for improving traffic network systems in the Hashemite Kingdom of Jordan, its capital Amman, and other regions globally. These studies primarily aim to optimize traffic networks for better control and flow management. Naveed *et al.* [1] demonstrated a system that used a support vector machine-based linear regression model. This model enables independent decision-making and reinforces the decision-making procedures within smart city wayside infrastructure. Boukerche *et al.* [2] presented traffic flow prediction procedures using statistical and machine learning models. Arabiat *et al.* [3] employed historical datasets to forecast traffic congestion. They focused on classifying the traffic congestion that occurs in the city of Amman, in particular, the Northbound Street on the 8th Circle, utilizing multiple machine learning models. The 8th Circle connects four main streets: Westbound, Northbound, Eastbound, and Southbound. Owing to the persistent jams in these areas, the northbound has been selected to be able to make a forecast about the congestion of traffic in that section. Traffic at the Northbound street is connected with 8th Circle, thus it is a reasonable objective to implement regression for the prediction of traffic congestion using four classifiers, including The Logistic Regression, Decision Tree, Random Forest, and multilayer perceptron. Four trials were conducted with the aid of the Waikato Environment for Knowledge Analysis software program, which aimed to identify the most suitable classifiers to tackle northbound traffic congestion. The selected classifiers were assessed using their F-measure, sensitivity, precision, and accuracy metrics. The findings from all experiments carried out point out that logistic regression is always the most effective tool in forecasting future traffic congestion. Jawarneh and Tawil [4] highlighted Amman's multifaceted causes of traffic congestion, linking them to land-use distribution and urban planning challenges, and provided recommendations for sustainable urban development. Laanaoui *et al.* [5] presented an innovative approach to urban traffic management, integrating real-time data processing and machine learning within Vehicular Ad Hoc Networks. The study addresses the challenges of increased vehicular density, data complexity, and inefficient traffic systems in urban environments, proposing a comprehensive solution aimed at improving traffic flow, reducing congestion, and enhancing safety. Salloum *et al.* [6] focused on enhancing network intrusion detection systems (NIDS) for intelligent vehicle systems by leveraging deep learning and synthetic data generation. Recognizing the limitations of traditional NIDS, which rely on static and often outdated datasets, the researchers utilized Scapy, a Python-based network manipulation tool, to create a realistic and extensive dataset of 100,000 network flows. These flows encompass a diverse range of benign, malicious, and outlier traffic patterns, mimicking real-world network scenarios. The generated dataset was analyzed using a Convolutional Neural Network, designed specifically for detecting network intrusions. Sathiyaraj *et al.* [7] presented a novel framework to address the persistent issue of urban traffic congestion, particularly tailored for Indian metropolitan cities. The proposed smart traffic prediction and congestion avoidance system incorporates a three-phase methodology: Traffic identification, prediction, and congestion avoidance. Leveraging a combination of Poisson distribution and genetic algorithms, the framework predicts traffic flow, identifies congestion patterns, and offers optimized solutions to mitigate delays and enhance fuel efficiency. The system begins by collecting real-time traffic data through sensors, which provide key parameters such as vehicle speed, arrival times, and volume. The Poisson distribution is employed to predict vehicle arrival rates, categorizing traffic conditions into low, medium, and heavy states. When congestion is predicted, the genetic algorithm is activated, optimizing traffic rerouting decisions based on a fitness function that evaluates energy efficiency, distance, and safety.

As the research presented in this article utilized traffic data from a street in Jordan, it is pertinent to explore related studies conducted in the Jordanian context. Al-Masaeid *et al.* [8] examined traffic volume prediction for rural (i.e., intercity) streets in Jordan, specifically the streets between Amman and Jerash and between Jerash and Irbid. This study evaluated three prediction methods: linear regression, trend analysis, and empirical Bayesian analysis. The dataset for this research included traffic volume records for the selected streets from 1996 to 2004, provided by the Ministry of Public Works and Housing of Jordan. The primary objective of the study was to estimate traffic volumes to support authorities in making informed decisions regarding street network planning and construction. Qaddoura and Bani Younes [9] utilized simulated traffic data from real streets in Jordan to predict traffic congestion levels in selected scenarios. The study focuses on streets located in Amman, including King Abdullah bin Al Hussein II Street, Queen Rania Al Abdullah Street, and Jordan Street. The machine learning methods applied in this research include linear regression, regression trees, and k-nearest neighbors (k-NN) regression. The features used to train the machine learning models include vehicle identity, acceleration, angle, distance, lane, position, signals, slope, speed, x-coordinate, and y-coordinate.

Jawarneh and Tawil [10] presented explores the application of machine learning techniques to predict pedestrian compliance at crosswalks in urban settings within Jordan, aiming to improve safety and traffic management. Four distinct machine learning algorithms were employed: artificial neural networks, support vector machines, decision trees, and random forests. The analysis revealed that pedestrian compliance is heavily influenced by local infrastructure and traffic conditions. Among the models, the random forest algorithm exhibited the highest accuracy and precision, making it the most reliable for predicting pedestrian behavior. The research concludes that integrating such predictive models into real-time traffic management systems could dynamically improve pedestrian safety and offer scalable solutions for broader urban challenges. Al Shafie *et al.* [11] investigated the application of auto-machine learning in predicting traffic accident severity in Jordan, addressing the challenges posed by road traffic accidents, which account for 1.35 million deaths globally each year and are the third leading cause of death in Jordan. The researchers utilized a dataset of 115,148 incidents and assessed multiple machine learning models, including decision trees, random forests, light gradient boosting machines, gradient boosting, extra trees, and bagging classifiers. The study employed metrics such as accuracy, balanced accuracy, recall, precision, and F1-score to evaluate the performance of these classifiers, both with and without hyperparameter tuning and down-sampling.

3. METHODOLOGY

3.1. Study Area and Problem Description

This study examines King Abdullah II Street in Amman, Jordan, focusing on the segment between the 8th Circle and Al-Sha'b Circle, which is also known as Business Park Circle. This corridor is one of the most heavily trafficked areas in Amman, especially during peak hours. The combination of high traffic density, the design of the circles, and the absence of effective traffic management systems results in considerable delays and frequent traffic congestion [12]. The analysis targets the morning rush hour (7:30 AM to 8:30 AM) since it reflects the time with the greatest traffic demand.

3.2. Data Collection

The Greater Amman Municipality supplied historical traffic data while field observations were used to collect traffic statistics. The data included:

- Traffic volume: The dataset measures the complete number of vehicles that traversed the study area in both northbound and southbound directions. 9,329 vehicles flowed from the 8th Circle towards Al-Sha'b Circle during peak hours, and 7,127 vehicles headed in the opposite direction during the same time period.
- Traffic composition: This data examines how vehicles are spread across different categories, such as passenger cars, along with buses and trucks. The traffic consisted of passenger cars, which represented 85% of total vehicles, and buses and trucks which together made up 15%.
- Traffic speed: The evaluation of vehicle speeds during peak and off-peak periods revealed typical traffic speeds. Average vehicle speeds fell to 10 km/h during peak hours but rose to 40 km/h when traffic was lighter.
- Delay times: The total amount of time vehicles waited at intersections and circles was averaged. Before starting the metering strategy, vehicles experienced an average delay time of 120 s at the 8th Circle and 110 s at Al-Sha'b Circle.

The Aimsun Next simulation model used collected data to accurately replicate actual traffic conditions [13].

3.3. Traffic Simulation Model

Aimsun Next served as the microscopic traffic simulation software, which facilitated detailed modeling of traffic flow alongside vehicle behavior and intersection dynamics for the simulation [14]. The simulation model included the following components:

- Network geometry: A complete model of the road network between the 8th Circle and Al-Sha'b Circle was developed with details about lanes, intersections, and traffic signals. The network segmentation process involved assigning specific traffic characteristics to each segment according to field data.
- Traffic demand: Fig. 1 shows a traffic demand matrix that recorded vehicle movement between centroids during the study period. Real traffic counts, along with assumptions for areas outside the study zone helped create the traffic matrix. The assumptions for traffic demand were obtained using Aimsun Next as the microscopic traffic simulation software.
- Vehicle behavior: Researchers conducted field observations to fine-tune the acceleration and deceleration parameters, along with lane-changing behavior and driver reaction times. The field observations revealed that drivers typically reacted in 1.5 s, and passenger cars could accelerate up to 2.5 m/s squared.
- Traffic control: The model included the current traffic signal timings and priority rules that govern the circles. Signal timings were modified to match real-world conditions observed during peak traffic hours.

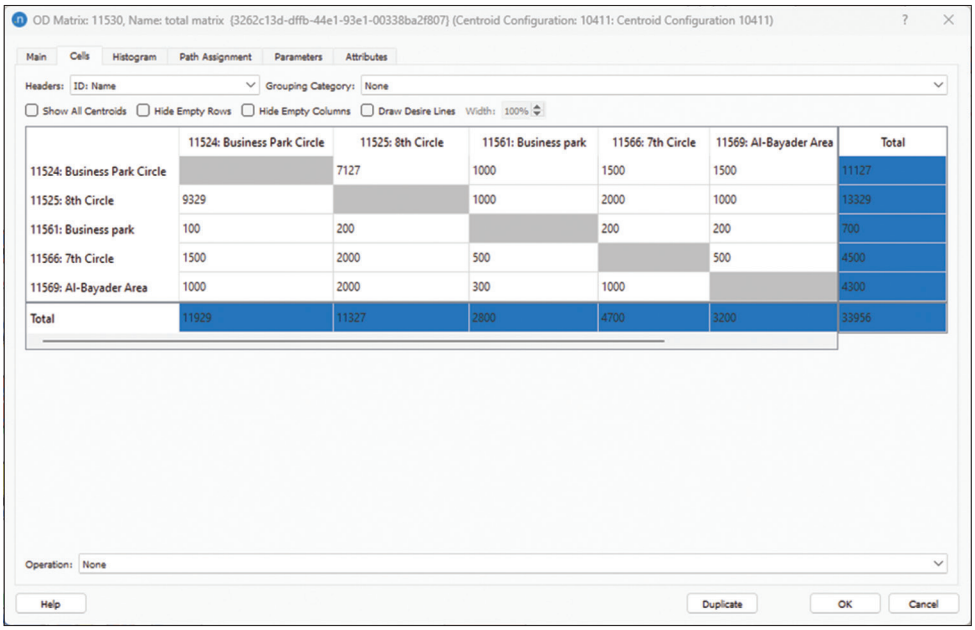
Fig. 1 shows the number of vehicles passing between the allocated centroids in the study area. The real numbers are between the 8th Circle and Al-Sha'b Circle (Business Park Circle), while the other are assumptions.

3.4. Traffic Metering Strategy

The simulation implemented a traffic metering strategy to deal with the traffic congestion. The traffic control plan required setting up metering points at every entrance to both the 8th Circle and Al-Sha'b Circle. All metering points required vehicles to wait for 5 s before granting access to the circle. Figs. 2-4 display examples of metering points positioned at key locations, which are:

- The route from Al-Bayader to the 8th Circle is illustrated in Fig. 2.
- The route from 7th Circle to 8th Circle is illustrated in Fig. 3.
- The route from Al-Korsi to Al-Sha'b Circle is illustrated in Fig. 4.

The metering system was created to control vehicle entry into the circles, which helps to stop congestion from forming while enhancing traffic movement [15]. Preliminary simulations demonstrated that a delay value of 5 s would effectively reduce congestion while avoiding excessive wait times for drivers.



	11524: Business Park Circle	11525: 8th Circle	11561: Business park	11566: 7th Circle	11569: Al-Bayader Area	Total
11524: Business Park Circle		7127	1000	1500	1500	11127
11525: 8th Circle	9329		1000	2000	1000	13329
11561: Business park	100	200		200	200	700
11566: 7th Circle	1500	2000	500		500	4500
11569: Al-Bayader Area	1000	2000	300	1000		4300
Total	11929	11327	2800	4700	3200	33956

Fig. 1. Origin-destination (OD) matrix.

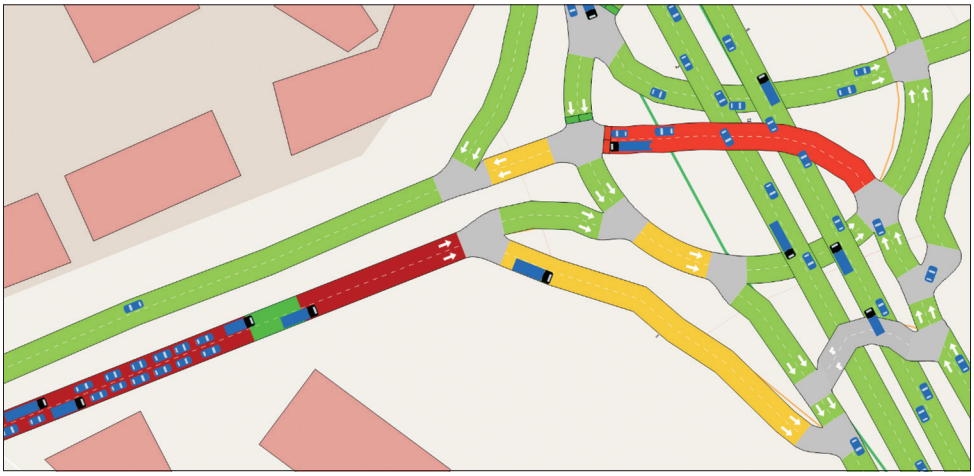


Fig. 2. Sample metering point on the path from Al-Bayader area to 8th circle.

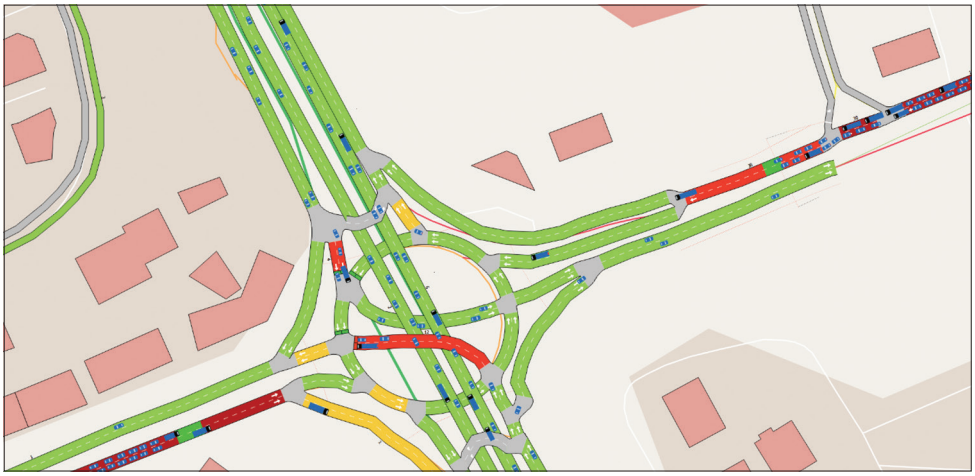


Fig. 3. Sample metering point on the path from the 7th Circle to 8th circle.

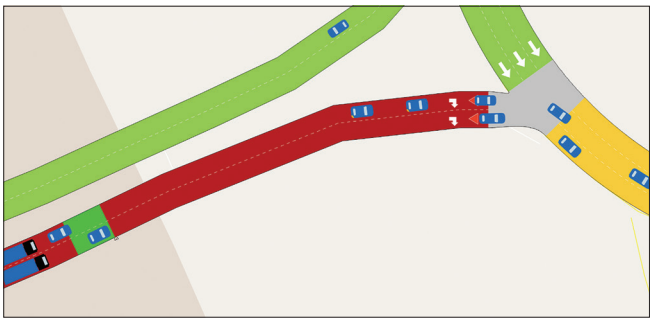


Fig. 4. Sample metering point along the path from Al-Korsi to Al-Sha'b circle.

3.5. Simulation Scenarios

- Baseline scenario: Under this scenario, the existing traffic conditions were assumed without any metering points. For peak traffic conditions, the simulation was performed for 1 h only (from 7:30 AM to 8:30 AM).
- Metering scenario: This scenario included the metering points in the 8th Circle and Al-Sha'b Circle The metering strategy effectiveness was evaluated running the simulation under the same conditions as the baseline scenario.

3.6. Evaluation Metrics

The effectiveness of the metering strategy was evaluated using the following metrics:

- Average delay time: The average time vehicles spent waiting in traffic at key intersections and circles [16].
- Queue length: The length of vehicle queues at the entry points to the circles [17].
- Travel time: The total time taken by vehicles to traverse the study area [18].
- Traffic flow: The number of vehicles passing through the study area during the simulation period [19].
- Congestion levels: The reduction in traffic jams and overall improvement in traffic flow [20].

Meters per second squared (m/s^2) is the unit of acceleration. It demonstrates how a vehicle's velocity changes over time. For example, if a vehicle accelerates at 2 m/s^2, its speed increases by 2 m/s every second.

4. RESULTS AND DISCUSSION

The baseline scenario showed extreme traffic congestion at both the 8th Circle and Al-Sha'b Circle. The data from the simulations indicate extensive vehicle queues and notable delays throughout the peak hour period (Figs. 5-9). For example:

- 8th circle: Traffic measurements indicate that vehicles experienced an average waiting time of 120 s while queues reached lengths of 200 m according to Fig. 5.
- Al-Sha'b circle: The typical wait time reached around 110 s while vehicle lines extended 180 m at Al-Sha'b Circle as shown in Fig. 6.
- King Abdullah II Street: Traffic patterns along King Abdullah II Street exhibited irregular flow dynamics, which resulted in repeated stop-and-go movements and extended delays as displayed in Figs. 7-9.

These figures conclude that the findings demonstrate that the current traffic management methods are ineffective and necessitate actions to alleviate congestion.

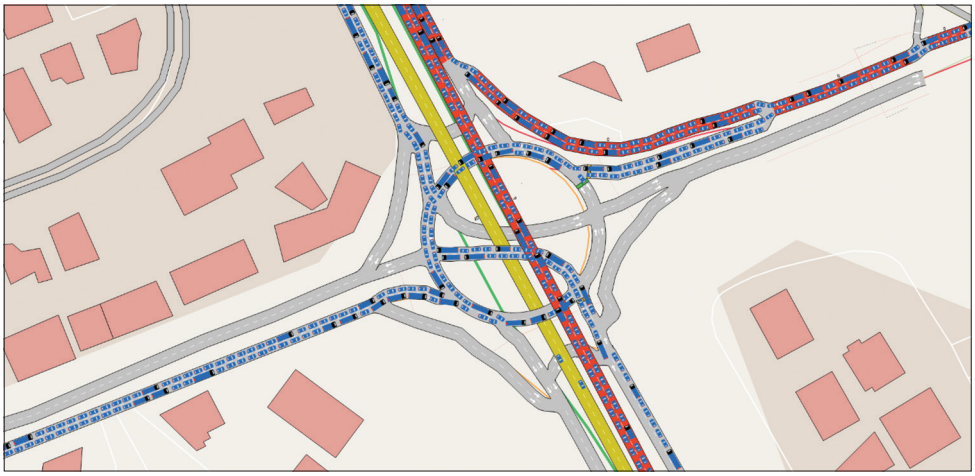


Fig. 5. Traffic conditions at the 8th circle before metering.

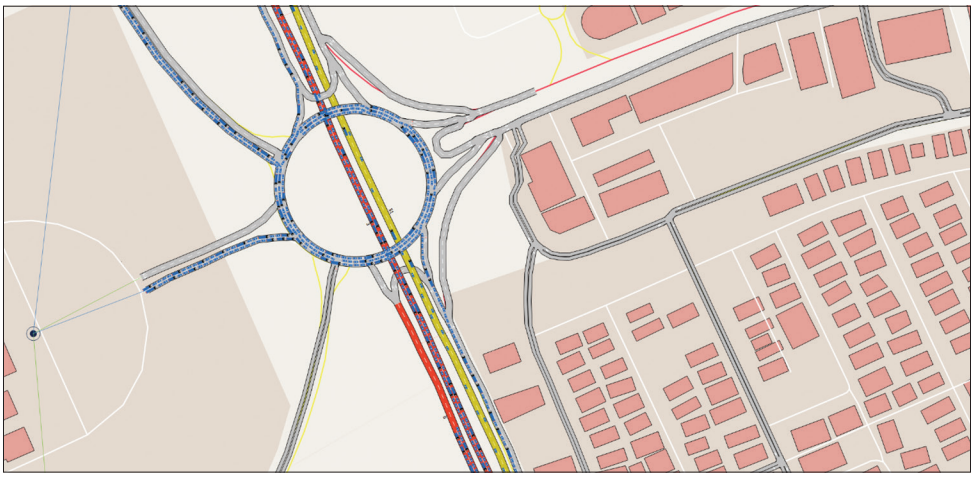


Fig. 6. Traffic conditions at Al-Sha'b circle before metering.

Fig. 7. Traffic conditions at King Abdullah II Street before metering (Morning Rush Hour).

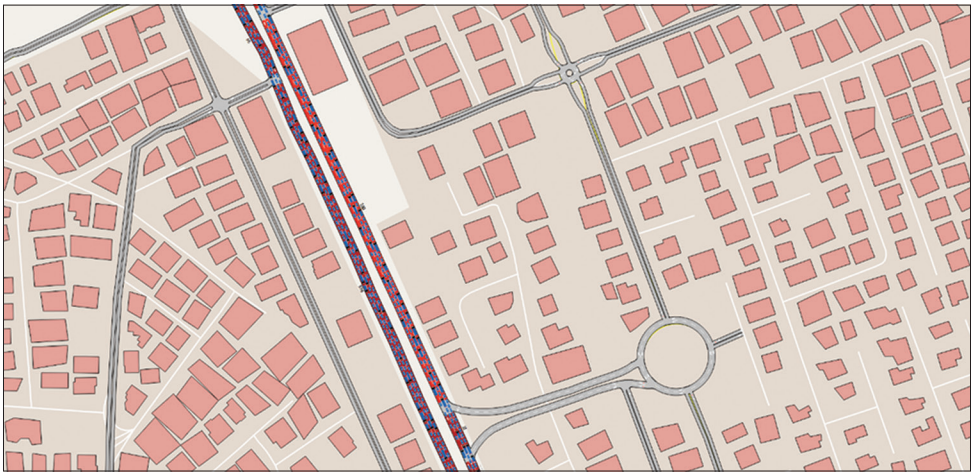


Fig. 8. Traffic conditions at King Abdullah II Street before metering (Midday).

Fig. 9. Traffic conditions at King Abdullah II Street before metering (Evening Peak Hour).

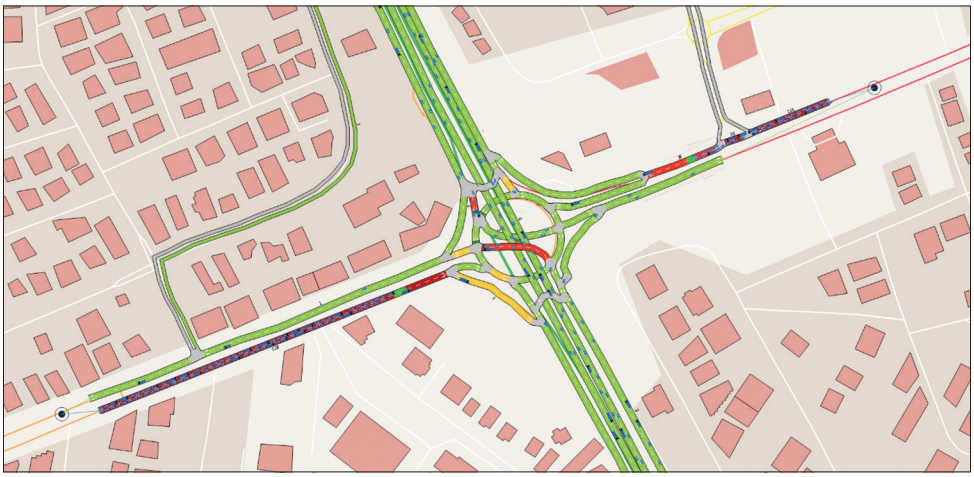


Fig. 10. Traffic conditions at 8th Circle after metering.

Traffic conditions underwent substantial enhancements after the metering strategy was put into action. The information presented in Figs. 10-14 illustrates the following outcomes.

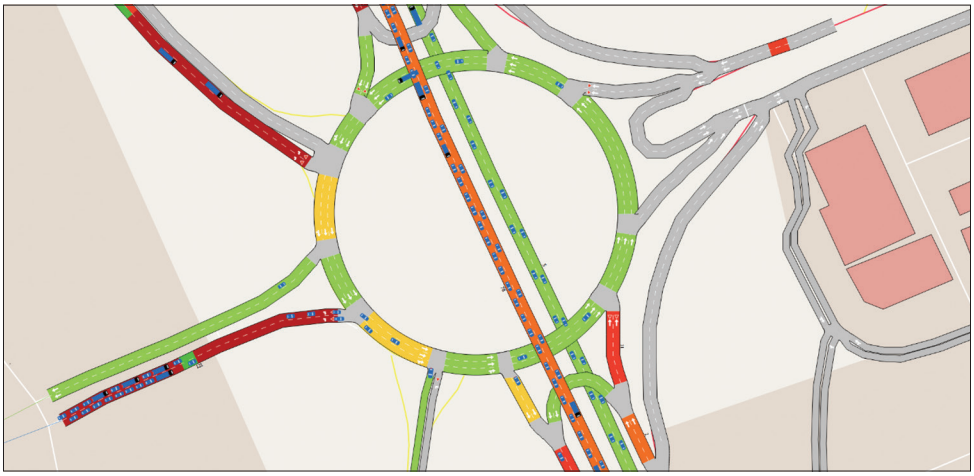


Fig. 11. Traffic conditions at Al-Sha'b Circle after metering.

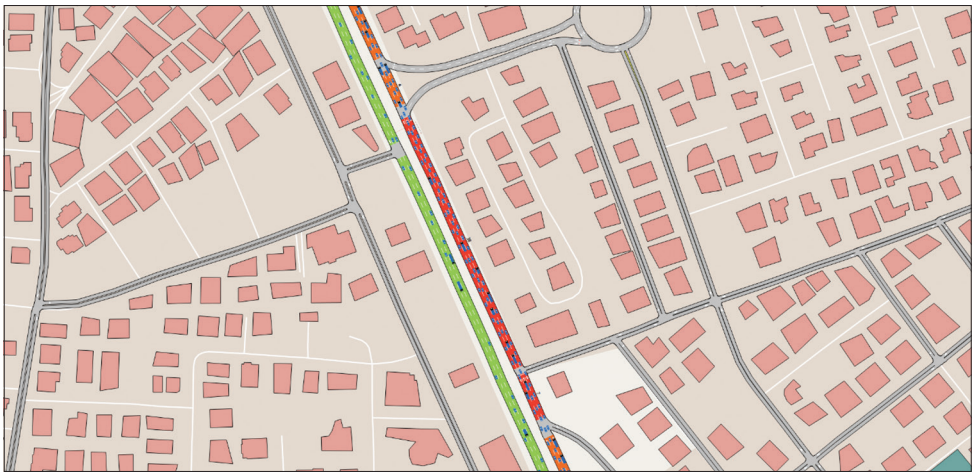


Fig. 12. Traffic flow at the northern segment of King Abdullah II Street after metering.

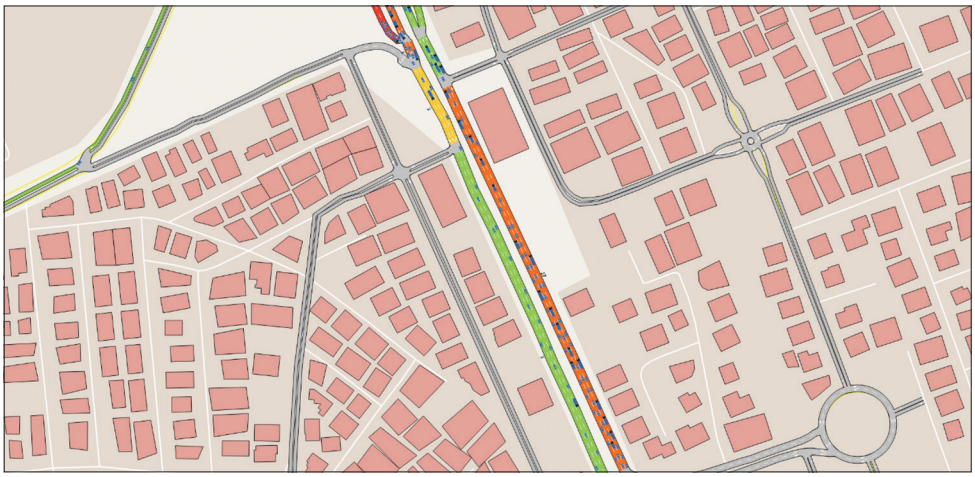


Fig. 13. Improved intersection throughput at King Abdullah II Street after metering.

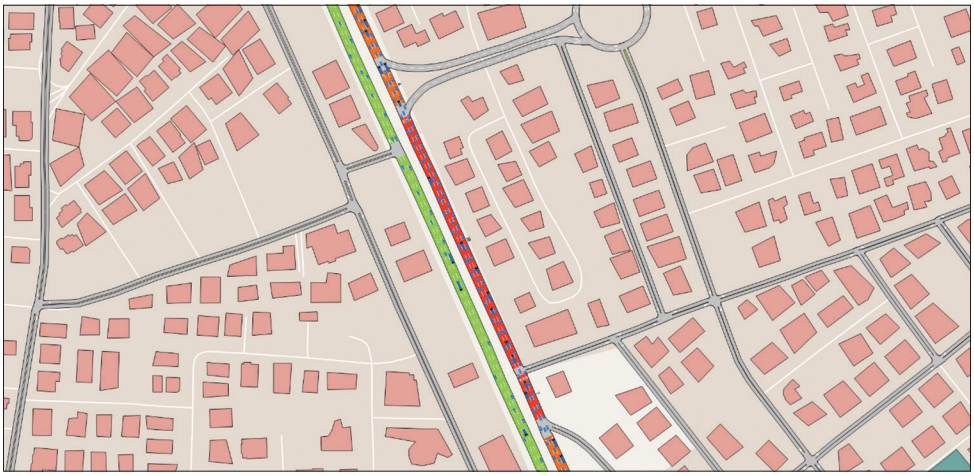


Fig. 14. Reduced congestion in the southern corridor of King Abdullah II Street after metering.

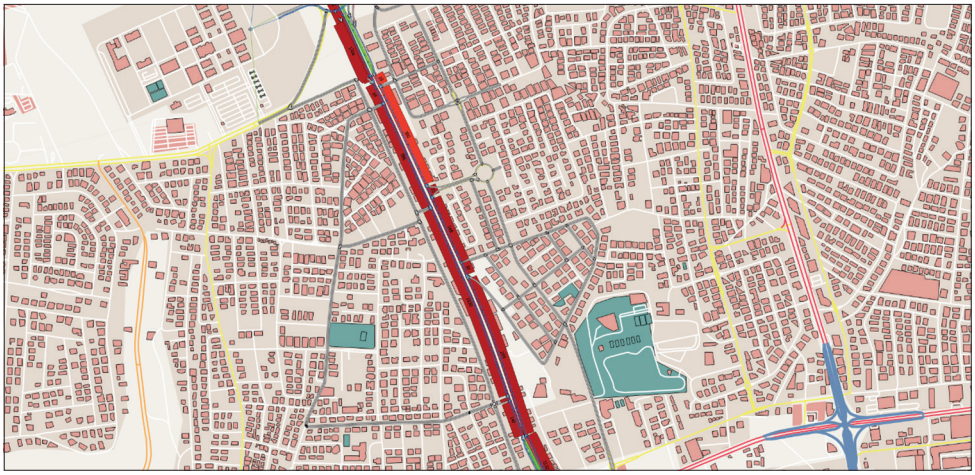


Fig. 15. Wide view of King Abdullah II Street before metering.

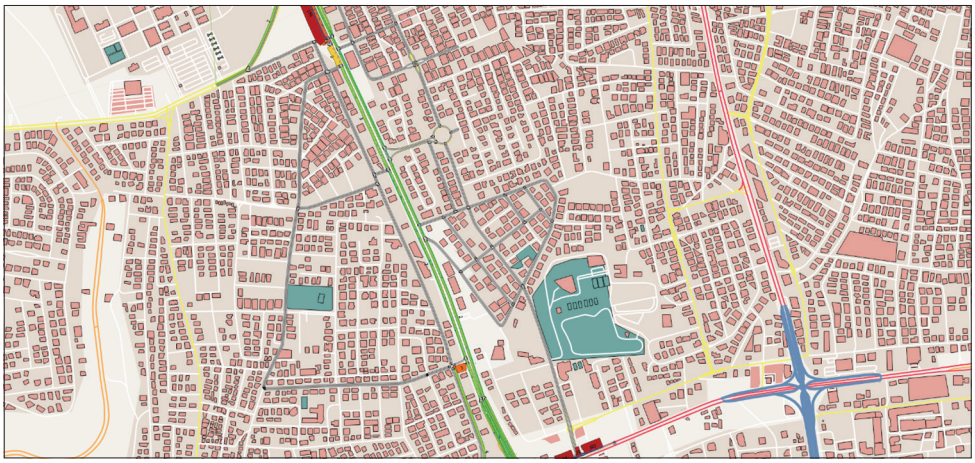


Fig. 16. Wide view of King Abdullah II Street after metering.

- 8th circle: Metering reduced the average delay time to 60 s from 120 s and the queue length diminished to 100 m as shown in Fig. 10.
- Al-Sha'b circle: The metering strategy reduced the average delay time from 110 s to 50 s while shortening the queue length to 80 m as shown in Fig. 11.
- King Abdullah II Street: The overall traffic pattern experienced greater stability through fewer congestion occurrences and smoother movement between intersections, as shown in Figs. 12-14.

Figs. 15 and 16 demonstrate how traffic conditions along King Abdullah II Street significantly improved. The metering strategy managed vehicle entry into the circles successfully, as shown by the decreased queue lengths and delay times, which stopped congestion from developing.

The simulation summary, which includes Figs. 17-20 demonstrates the delay time comparison before and after the metering strategy was applied. Thus, the summary shows that average delay times noticeably decreased in the study area. For example:

- 8th circle: The study area saw average delay times cut by half, as they reduced from 120 s to 60 s according to Figs. 17 and 18.
- Al-Sha'b circle: The metering strategy implementation led to a 54.5% reduction in delay time from 110 s to 50 s according to Figs. 19 and 20.

These results confirm that the metering strategy effectively reduced traffic congestion and improved traffic flow.

The study findings show that traffic metering points create a substantial reduction in congestion for high-traffic areas. The metering approach reduced congestion and maintained smooth traffic flow through key intersections by briefly holding vehicles for 5 s before they entered. The metering strategy works best on roads that experience high traffic

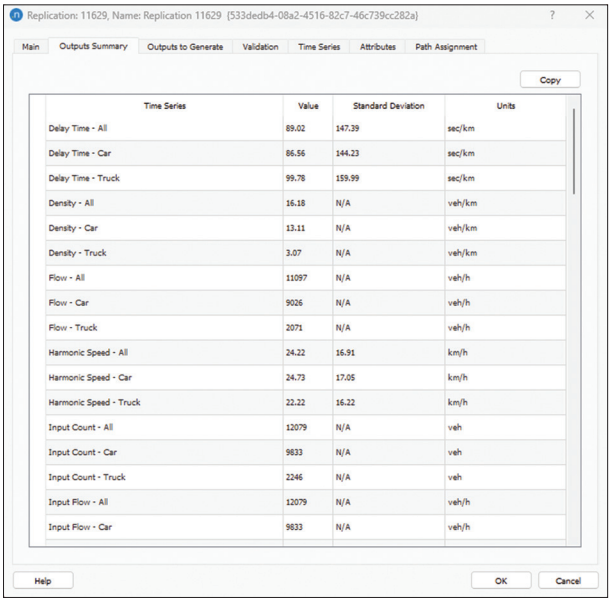


Fig. 17. Simulation summary of cars delays in King Abdullah II Street after utilizing metering.

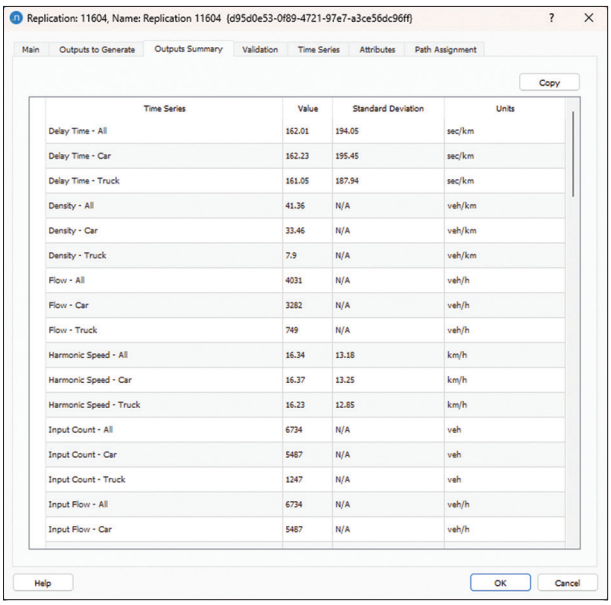


Fig. 18. Simulation summary of car delays in King Abdullah II Street before utilizing metering.

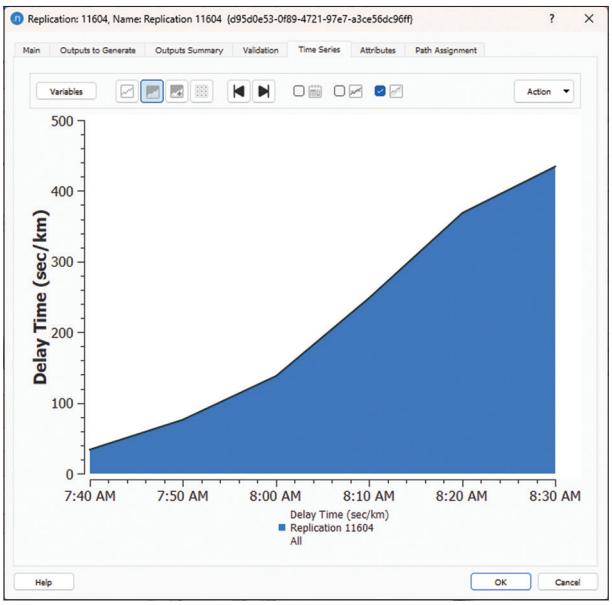


Fig. 19. Summary of simulation, average delay time before utilizing metering.

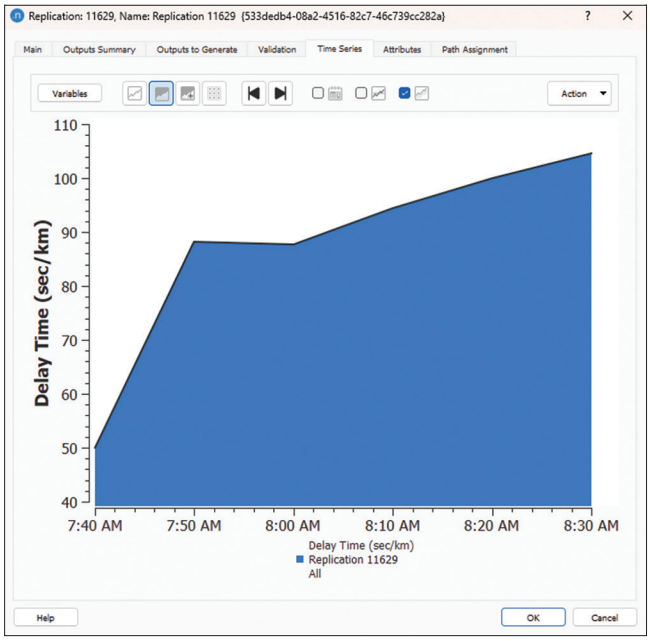


Fig. 20. Summary of simulation, average delay time after utilizing metering.

during rush hour, like King Abdullah II Street in Amman. By balancing vehicle entry into circles, this metering strategy achieves success because it reduces the chances of traffic jams forming. The effectiveness of this strategy varies based on driver behavior patterns, as well as road conditions and the availability of alternative routes.

5. CONCLUSION

This research examined the impact of a traffic metering strategy on reducing congestion along King Abdullah II Street in Amman, from the 8th Circle to Al-Sha'b Circle (Business Park Circle). Through field observations and historical traffic data supported by microscopic simulation with Aimsun Next, the study analyzed the traffic conditions before and after metering point implementation. Traffic improvements emerged through metering strategy application, which notably decreased delays while enhancing flow and reducing congestion, specifically during peak hours. Before introducing metering systems, traffic analysis showed extreme congestion at the 8th circle with 120-s average delays and at Al-Sha'b Circle with 110-s delays; both locations experienced vehicle queues that reached 200 m and 180 m, respectively. The inconsistent traffic movement that featured regular stop-and-go patterns demonstrated the shortcomings of the current traffic control system. The metering approach manages vehicle entry to essential intersections, which stops congestion from forming and shortens delays while improving traffic flow efficiency. Future research must evaluate how metering strategies affect traffic flow and driver behavior alongside road safety over long periods to determine their lasting viability. In addition, the application of the metering strategies could spread to other congested areas in Amman. Moreover, smart traffic management systems benefit from the integration of metering techniques, which enable real-time traffic monitoring and dynamic traffic signal adjustments.

REFERENCES

[1] S. A. Naveed, U. Farooq, M. A. Raza, Z. U. Rehman, M. Saleem, and T. M. Ghazal. "Enhancing traffic flow and congestion management in smart cities utilizing SVM-based linear regression approach". International Journal of Advanced and Applied Sciences, vol. 11, no. 10, pp. 166-175, 2024.

[2] A. Boukerche, Y. Tao and P. Sun. "Artificial intelligence-based vehicular traffic flow prediction methods for supporting intelligent transportation systems. *Computer Networks*, vol. 182, p. 107484, 2020.

[3] A. Arabiat, M. Hassan and Al Momani, O. "Traffic congestion prediction using machine learning: Amman city case study". *Proceedings of SPIE*, vol. 13188, p. 1318806, 2024.

[4] H. Jawarneh and M. Tawil. "Exploring the interrelation between traffic congestion and land uses in Amman: Challenges and potential". *Journal for Re-Attach Therapy and Developmental Diversities*, vol. 6, no. 9s, pp. 962-978, 2023.

[5] M. D. Laanaoui, M. Lachgar, H. Mohamed, H. Hrimech, S. G. Villar and I. Ashraf. "Enhancing urban traffic management through real-time anomaly detection and load balancing". *IEEE Access*, vol. 12, pp. 63683-63698, 2024.

[6] S. A. Salloum, T. Gaber, M. A. Almaiah, R. Shehab, R. Al-Ali and T. H. H. Aldahyani. "Adoption of a deep learning approach using realistic synthetic data for enhancing network intrusion detection in intelligent vehicle systems". *International Journal of Data and Network Science*, vol. 9, no. 1, pp. 77-86, 2025.

[7] R. Sathiyaraj, A. Bharathi, S. Khan, T. Kiren, I. U. Khan and M. Fayaz. "A genetic predictive model approach for smart traffic prediction and congestion avoidance for urban transportation". *Wireless Communications and Mobile Computing*, vol. 2022, p. 5938411, 2022.

[8] H. R. Al-Masaeid and N. J. Al-Omoush. "Trafc volume forecasting for rural roads in Jordan". *Jordan Journal of Civil Engineering*, vol. 8, no. 3, pp. 319-331, 2014.

[9] R. Qaddoura and M. Bani Younes. "Temporal prediction of traffic characteristics on real road scenarios in Amman". *Journal of Ambient Intelligence and Humanized Computing*, vol. 14, no. 7, pp. 9751-9766, 2023.

[10] M. M. Taamneh, A. H. Alomari and S. M. Taamneh. "Using machine learning to predict pedestrian compliance at crosswalks in Jordan". *Applied Sciences*, vol. 14, no. 11, p. 4945, 2024.

[11] R. Al Shafie, K. H. Jadaan, S. H. El-Badawy, S. Shwaly and U. Shahdah. "Exploring the utility of auto-machine learning in predicting traffic accident severity in Jordan". *Jordan Journal of Civil Engineering*, vol. 18, no. 4, pp. 583-594, 2024.

[12] Aimsun Next User Manual. "*Aimsun Next Simulation Software*". 2023. Available from: https://www.aimsun.com/aimsun-next-

downloads [Last accessed on 2025 Feb 03].

[13] H. R. Al-Masaeid and T. I. Al-Suleiman. “Traffic flow analysis in Urban areas: A case study of Amman, Jordan”. *International Journal of Civil Engineering*, vol. 19, no. 4, pp. 567-578, 2021.

[14] B. S. Kerner. “*The Physics of Traffic: Empirical Freeway Pattern Features, Engineering Applications, and Theory*”. Springer, Berlin, Germany, 2024.

[15] J. Barceló. “*Fundamentals of Traffic Simulation*”. Springer, Berlin, Germany, 2010.

[16] C. F. Daganzo. “*Fundamentals of Transportation and Traffic Operations*”. Pergamon, United Kingdom, 1997.

[17] E. Brockfeld, R. D. Kühne and P. Wagner. “Calibration and validation of microscopic traffic flow models”. *Transportation Research Record: Journal of the Transportation Research Board*, vol. 1876, pp. 62-70, 2004.

[18] M. Treiber and A. Kesting. “*Traffic Flow Dynamics: Data, Models and Simulation*”. Springer, Berlin, Germany, 2013.

[19] R. W. Rothery. “*Car Following Models, Traffic Flow Theory*”. Transportation Research Board, Washington, DC, 1992.

[20] T. V. Mathew. “*Transportation Systems Engineering*”. IIT Bombay, 2014. Available from: https://www.transportation_engg_-_tom_v_mathew.pdf [Last accessed on 2025 Mar 01].

ORIGINAL RESEARCH ARTICLE UHD JOURNAL OF SCIENCE AND TECHNOLOGY

Surgery Versus Flexible Endoscopic Rubber Band Ligation for Grade 2 and 3 Internal Hemorrhoids

Diyaree Nihad Ismael

Senior General Surgeon, Department of Surgery, Teaching Hospital, Sulaimanyah, Iraq

ABSTRACT

Surgery has traditionally been the primary treatment for symptomatic internal hemorrhoids. However, office-based interventions such as rubber band ligation (RBL) are increasingly used for Grades 1–3 hemorrhoids. Flexible endoscopic RBL offers a minimally invasive alternative, whereas surgery remains standard for Grade 4. To compare the effectiveness of flexible endoscopic RBL versus surgical hemorrhoidectomy in managing symptomatic Grades 1–3 internal hemorrhoids, focusing on bleeding control, pain, recovery time, and recurrence. A comparative study of 55 patients treated with flexible endoscopic RBL (using Olympus kits) and 55 matched patients undergoing conventional excisional hemorrhoidectomy (open technique). Patients choose their treatment after counseling. Outcomes were assessed over 1 year, with follow-up at 1 week, 3, 6, and 12 months. Pain was measured using a Visual Analog Scale (≥4 defined significant pain). Statistical analysis used a statistical package for the social sciences v26 (t-tests for continuous variables, Chi-square for categorical; $P < 0.05$ significant). Both groups showed comparable efficacy: Bleeding control (95% vs. 93%), mucosal prolapse resolution (96% vs. 97%), and 1-year recurrence (30% vs. 29%). RBL had superior post-procedural outcomes: Lower pain (10% vs. 90%), fewer work absences (5% vs. 95%), and no bed-boundness (0% vs. 100%; all $P < 0.05$). Flexible endoscopic RBL is as effective as surgery for Grades 1–3 hemorrhoids but significantly reduces pain, recovery time, and work absenteeism. RBL should be considered a first-line option for eligible patients.

Index Terms: Rubber Band Ligation, Internal Piles, Hemorrhoidectomy

1. INTRODUCTION

Hemorrhoidal disease is the leading condition affecting the rectum and anal canal, with a global prevalence ranging from 2.9% to 27.9%, where more than 4% of cases are symptomatic [1], [2]. Approximately one-third of these individuals consult physicians for guidance. The age distribution follows a Gaussian curve, peaking between 45 and 65 years and tapering off after 65 years [1]. Men are more commonly affected than women [3]. Hemorrhoids are believed to result from the downward displacement of vascular cushions caused by a disrobing of the supporting suspensory Treitz's muscle [3]. Several factors, such as a low-fiber diet, prolonged straining, constipation, diarrhea, and hard stools, can trigger hemorrhoidal symptoms [3]. Symptoms may include rectal bleeding, prolapse of the hemorrhoidal cushions, pain due to thrombosis, itching-related discomfort, mucus discharge, and fluid incontinence. Internal hemorrhoids are categorized into four degrees based on the extent of prolapse: First-degree (non-prolapsing), second-degree (prolapsing during straining but reducing spontaneously), third-degree (prolapsing during straining and requiring manual reduction), and fourth-degree (permanently prolapsed) [4]. Notably, the severity of symptoms does not

Access this article online

DOI:10.21928/uhdjst.v9n1y2025.pp181-184

E-ISSN: 2521-4217

P-ISSN: 2521-4209

Corresponding author's e-mail: Diyaree Nihad Ismael, Senior General Surgeon, Department of Surgery, Teaching Hospital, Sulaimanyah, Iraq. E-mail: Nihaddiyaree@gmail.com

Received: 16-12-2024 Accepted: 23-04-2025 Published: 15-06-2025

always align with the degree of hemorrhoids. Treatment options for symptomatic hemorrhoids have evolved over time, encompassing conservative medical management, non-surgical methods, and various surgical techniques, including stapled hemorrhoidopexy. Medical interventions such as recommending a high-fiber diet and bulk-forming agents can effectively prevent constipation and the associated complications of hemorrhoids [5]. In addition, numerous commercial ointments are available for symptomatic relief, although evidence supporting their efficacy is limited [6]. Phlebotonics, such as flavonoids, are also used. Non-surgical treatments include rubber band ligation (RBL), injection sclerotherapy, cryotherapy, infrared coagulation, laser therapy, and diathermy coagulation – all of which can be performed as outpatient procedures without anesthesia. These non-surgical methods are considered the gold standard for managing grade one to three (grade I-III) hemorrhoids [7]. Among all the non-surgical procedures, RBL stands out as the most effective in terms of patient compliance, long-term success, and minimal side effects [7]. RBL is a straightforward, quick, and cost-effective outpatient procedure first introduced by Blaisdell and later refined by Barron [7]. The procedure involves applying rubber bands to an insensitive area just above the dentate line, with up to three bands applied in one session, which can be safely repeated after 4–6 weeks. Various techniques, including endoscopic ligation, are used for band application, but the suction method is the most common. Studies report success rates for RBL ranging from 69% to 94% [8]. RBL has a low complication rate of <2%. Possible complications include vasovagal syncope, anal pain, minor bleeding, chronic ulceration, priapism, difficulty urinating, thrombosis of external hemorrhoids, and, in rare cases, severe complications such as massive bleeding or pelvic sepsis [8]. If conservative measures fail to alleviate symptoms, patients are referred to a surgeon for operative management. Surgical treatment is indicated for cases with significant external components, hypertrophied papillae, associated fissures, extensive thrombosis, or recurring symptoms after repeated RBL.

Post-hemorrhoidectomy pain is the most common challenge associated with surgical interventions. Early complications include urinary retention (20.1%), bleeding (secondary or reactionary) (2.4–6%), and subcutaneous abscesses (0.5%). Long-term complications can include anal fissure (1–2.6%), anal stenosis (1%), incontinence (0.4%), fistula formation (0.5%), and hemorrhoid recurrence [9]–[11]. This study aimed to compare the effectiveness, post-operative outcomes, and recurrence rates of flexible endoscopic RBL versus surgical hemorrhoidectomy in patients with symptomatic Grade 2 and 3 internal hemorrhoids, with a focus on bleeding control, pain, recovery time, and long-term recurrence.

2. METHODS

A cross-sectional comparative study was conducted from January 2021 to March 2024 in Sulaymaniyah, Iraq, across public and private hospitals of 55 patients with symptomatic internal hemorrhoids of Grades 1–3 treated with flexible endoscopic RBL using the rubber band kits of Olympus company containing six rubber bands pre-loaded on plastic cups. The outcome of these patients was compared with another 55 patients, matched for most of their characteristics, and treated with conventional excisional hemorrhoidectomy (open surgical technique). The patients were free to choose one of the two above options after a good explanation of the characteristics, effectiveness, possible adverse events, and recurrence rates of both methods, depending on available literature. The duration of follow-up of both patient groups was 1 year. Patients were assessed at 1 week, 3 months, 6 months, and 12 months post-intervention for bleeding, prolapse, pain, and recurrence through clinical examination. Post-operative pain was evaluated using a Visual Analog Scale (VAS) ranging from 0 (no pain) to 10 (worst imaginable pain). Pain scores were recorded at 6-, 24-, and 48-h post-intervention. Patients reporting a VAS score ≥4 were categorized as having significant pain (10% in RBL vs. 90% in surgery groups, respectively). We used a flexible gastroscope of Olympus type 160 and a rubber band Ligator kit of Olympus type of the same used for banding esophageal varices. Statistical analysis used *Statistical Package for the Social Sciences (SPSS) v26;* continuous variables were analyzed with Student's t-test, and categorical outcomes with Chi-square tests. Significance was set at $P<0.05$.

3. RESULTS

Characteristics of both groups of patients showed no

TABLE 1: Characteristics of both patient groups: band ligation group and surgery group

Characteristics	Band ligation group	Surgery group	*P*-value
Age	38.2±8.5	39.1±9.2	0.62
Sex M/F	30/25	32/23	
Smoking	15/55	17/55	
Obesity (BMI ≥30)	5/55	10/55	
Benign prostatic hyperplasia	1/55	1/55	
Constipation	40/55	41/55	

BMI: Body mass index

statistically significant differences, as shown in Table 1.

Table 2 shows the difference between the two intervention groups regarding the outcome of the intervention, showing no statistically significant difference in intervention outcomes, namely bleeding (95% vs. 93%), symptomatic mucosal protrusion (96% vs. 97%), and 1-year recurrence rate (30% vs. 29%).

Table 3 shows a statistically significant difference between the two intervention groups in favor of band ligation including post-operative pain (10% vs. 90%), work absence (5% vs. 95%), and 1 week bed-boundness (0.00% vs. 100%).

Pain percentages reflect patients with VAS ≥4 (see methods for details).

4. DISCUSSION

There has been growing interest in the use of non-invasive, non-surgical management of symptomatic hemorrhoids as in other conditions in which surgery was the main approach because surgical hemorrhoidectomy typically requires anesthesia (general, spinal, or local), and unlike office-based RBL, it involves tissue excision, which increases post-operative pain and recovery time. Many office-based hemorrhoidal interventions use specific office instruments such as banding by rigid instruments and laser therapy by special probes [1]–[3]. The use of flexible endoscopy to do band ligation for internal piles provides flexibility to see and ligate all the hemorrhoid columns. In addition, the flexible scope can do ligation in the retroflexed and forward approach [12]. We performed this study on two groups of patients, each including 55 patients with symptomatic internal hemorrhoids, to compare their effectiveness and adverse events.

TABLE 2: Difference in outcomes between the two intervention groups

Interventions outcomes	Band ligation(%)	Surgery(%)	*P*-value
Bleeding control	95	93	>0.05
Mucosal protrusion disappearance	96	97	
1-year recurrence	30	29	

TABLE 3: Rate of complication of both groups

Complications	Band ligation(%)	Surgery(%)	*P*-value
Pain	10	90	<0.05
Work absence	5	95	
1-week bed boundness	0.00	100	

There was no statistically significant difference in the patient characteristics of both groups, such as age, sex, smoking status, obesity, prostatic enlargement, and obesity, to negate any effects of the comparison results. The literature review shows the effect of these variables on the incidence of hemorrhoids [1]. There were no statistically significant differences in the effectiveness of both interventions in controlling the symptoms of internal piles of Grades 1–3, such as bleeding, mucosal prolapse, and 1-year recurrence. It shows that both interventions control similarly both bleeding and mucosal prolapse, and interestingly each intervention has the same recurrence rate of around one-quarter to one-third of patients. Other studies and reviews confirm our findings, especially the recurrence rate [1]–[3]. Recurrence rates of hemorrhoids after any intervention are common if the causative or aggravating factors are not controlled, such as constipation, smoking, and prostatic enlargement, which are most of the time difficult to control [5].

There were statistically significant differences between the two interventions regarding post-intervention pain, work absence, and 1-week bed-boundness in favor of band ligation. Other studies confirmed these results, especially pain, which rarely occurs with band ligation as they are painless unless some bands are misfired and ligated columns above the dentate line which is a very painful area [9], [10]. A Cochrane review states that RBL should be preserved for Grade 2 and then surgery if this fails [1].

5. CONCLUSION

Comparing flexible endoscopic RBL and surgery, there was no significant difference in controlling symptomatic Grades 1–3 internal hemorrhoids, namely bleeding, mucosal protrusion, and 1-year recurrence, while there were significant differences in post-interventional pain, work absence, and ambulance being in favor of band ligation.

6. RECOMMENDATIONS

We recommend flexible endoscopic RBL for symptomatic Grades 1–3 included after a good explanation of the available

approved information on the available options for managing them.

REFERENCES

[1] V. Shanmugam, A. Hakeem, K. L. Campbell, K. S. Rabindranath, R. J. C. Steele, M. A. Thaha, M. A. Loudon. "Rubber band ligation versus excisional hemorrhoidectomy for hemorrhoids". *Cochrane Database of Systematic Reviews*, vol. 20, no. 11. p. CD005034, 2011.

[2] U. Ali and A. S. Khan. "Rubber band ligation versus open hemorrhoidectomy: A study of 100 cases". *Journal of Postgraduate Medical Institute*, vol. 19, no. 3. p. 317-322, 2005.

[3] B. R. Davis, S. A. Lee-Kong, J. Migaly, D. L. Feingold and S. R. Steele. "The American society of colon and rectal surgeons clinical practice guidelines for the management of hemorrhoids". *Diseases of the Colon and Rectum,* vol. 61, no. 3, pp. 284-292. 2018.

[4] V. Lohsiriwat. "Hemorrhoids: From basic pathophysiology to clinical management". *World Journal of Gastroenterology*, vol. 18, no. 17, pp. 2009-2017, 2012.

[5] S. R. Brown. "Hemorrhoids: Diagnosis and current management." *Annals of the Royal College of Surgeons of England*, vol. 99, no. 1, pp. 8-14, 2017.

[6] D. F. Altomare and S. Giuratrabocchetta. "Conservative and surgical treatment of haemorrhoids". *Nature Reviews Gastroenterology and Hepatology*, vol. 10, no. 9, pp. 513-521, 2013.

[7] V. S. Iyer, I. Shrier and P. H. Gordon. "Long-term outcome of rubber band ligation for symptomatic primary and recurrent internal hemorrhoids". *Diseases of the Colon and Rectum*, vol. 47, no. 8, pp. 1364-1370.

[8] H. M. MacRae and R. S. McLeod. "Comparison of hemorrhoidal treatment modalities: A meta-analysis". *Diseases of the Colon and Rectum*, vol. 38, no. 7. pp. 687-694, 1995.

[9] A. M. El Nakeeb, A. A. Fikry, W. H. Omar, A. A. Fikry, W. H. Omar, E. M. Fouda, T. A. El Metwally, H. E. Ghazy, S. A. Badr, M. Y. Abu Elkhar, S. M. Elawady, H. H. Abd Elmoniam, W. W. Khafagy, M. M. Morshed, R. E. El Lithy and M. E. Farid. "Rubber band ligation for 750 cases of symptomatic hemorrhoids out of 2200 cases". *World Journal of Gastroenterology*, vol. 14, no. 42, pp. 6525-6530, 2008.

[10] P. Giordano, J. Overton, F. Madeddu, S. Zaman and G. Gravante. "Transanal hemorrhoidal dearterialization: A systematic review". *Diseases of the Colon and Rectum*, vol. 52, no. 9, pp. 1665-1671.

[11] S. Jayaraman, P. H. Colquhoun and R. A. Malthaner. "Stapled hemorrhoidopexy is associated with a higher long-term recurrence rate of internal hemorrhoids compared with conventional excisional hemorrhoid surgery". *Diseases of the Colon and Rectum*, vol. 50, no. 9, pp. 1297-1305, 2007.

[12] G. Gallo, J. Martellucci, A. Sturiale, G. Clerico, G. Milito, F. Marino, G. Cocorullo, P. Giordano, M. Mistrangelo and M. Trompetto. "Consensus statement of the Italian society of colorectal surgery (SICCR): Management and treatment of hemorrhoidal disease". *Techniques in Coloproctology*, vol. 24, no. 2, pp. 145-164, 2020.

p-ISSN 2521-4209
e-ISSN 2521-4217

جامعة التنمية البشرية
UNIVERSITY OF HUMAN DEVELOPMENT

UHD Journal of Science and Technology

A Scientific periodical issued by University of Human Developement

Vol.9 No.(1) June 2025

2025 2725

e.mail:jst@uhd.edu.iq