Photodiagnosis and Photodynamic Therapy 54 (2025) 104621

Contents lists available at ScienceDirect

Photodiagnosis and Photodynamic Therapy

journal homepage: www.elsevier.com/locate/pdpdt

Research paper

Enhanced glaucoma detection using U-Net and U-Net+ architectures using deep learning techniques

B.P. Pradeep kumar [a], Pramod K.B. Rangaiah [b], Robin Augustine [b,*]

[a] Department of Computer Science and Design, Atria Institute of Technology, Bengaluru 560024, India
[b] Microwaves in Medical Engineering Group, Division of Solid State Electronics, Department of Electrical Engineering, Uppsala University, Box 65, SE-751 03, Uppsala, Sweden

ARTICLE INFO

Keywords:
Glaucoma
Retinal intra-ocular region
Frame networks
U-NET layers
DRISHTI-GS
DRIONS-DB
HRF

ABSTRACT

This study compares multiple image processing and deep learning methods to demonstrate an enhanced approach to glaucoma diagnosis. The approach focuses on noise reduction using median filtering and optic disc segmentation utilizing the U-Net and U-Net+ architectures. Capsule Networks were utilized for feature extraction and Extreme Learning Machines (ELM) for diagnostic classification. Three datasets were evaluated, including DRISHTI-GS, DRIONS-DB, and HRF, utilizing important parameters such as accuracy, sensitivity, and specificity. The findings revealed that median filtering reduced noise by 97.88%, with a peak signal-to-noise ratio of 44.99. U-Net beat U-Net+ in optic disc in the process of segmentation with a Dice coefficient of 0.8557, a Jaccard index of 0.7307, and higher segmentation accuracy. The suggested model has great diagnostic accuracy, scoring 99% for DRISHTI-GS, 99.5% for DRIONS-DB, and 98.5% for HRF. These findings show that using deep learning approaches can increase glaucoma diagnosis accuracy and reliability, with important implications for healthcare applications and patient outcomes.

1. Introduction

The realm of ocular disorders, glaucoma emerges as a complex challenge where diagnostic intricacies and therapeutic evolution intersect with the domain of image processing [1]. This scientific discourse embarks on an exhaustive journey to expound the diverse taxonomy of glaucoma subtypes ranging from Open-Angle, Angle-Closure, to Normal Tension Glaucoma each distinguished by distinct pathophysiological mechanisms unraveling in the intricate weave of ocular images [2]. The interplay of multifaceted risk determinants, encompassing advanced age dynamics, hereditary predisposition unraveling through genetic imprints within images, elevated intraocular pressure depicted in tonometry curves, and ethnicity encoded variations etched within images, amalgamate to form the rich mosaic of glaucoma susceptibility, further enriched by quantitative image analyses [3].

Embedded within the clinical narrative is glaucoma's silent predilection a stealthy advancement that eludes early perception, surreptitiously eroding the visual expanse from periphery to center [4]. This latent progression accentuates the role of image processing in diagnosis and monitoring. With image processing tools like optic nerve head morphometry, retinal nerve fiber layer thickness assessments using Optical Coherence Tomography (OCT), and optic disc evaluations employing ophthalmoscopy, the diagnostic landscape is unveiled in unprecedented detail [5].

Navigating the therapeutic course, the symphony of image processing harmonizes with personalized care [6]. Pharmacological interventions, meticulously choreographed through images of intraocular pressure trends, encompass beta-adrenergic blockers, prostaglandin analogs, and carbonic anhydrase inhibitors. For images that exhibit defiance, lasers like selective laser trabeculoplasty manifest as guiding beacons in modulating aqueous dynamics [7]. Surgical interventions portrayed in the images trabeculectomy and drainage implants [8] stand as the last fortress when images depict unyielding pressure.

In this grand tapestry of image-driven glaucoma management, the resonance of preventive measures resounds [9]. Images captured and analyzed through regular ocular surveillance, especially for images of high-risk groups illuminate early detection and guide intervention strategies [10]. Concurrent health images, notably of diabetes mellitus and hypertension, demand meticulous evaluation, recognizing their role as influential actors in the image of glaucoma pathogenesis [11].

Mired in the art and science of image interpretation [12], the expertise of accomplished ophthalmologists intertwines with image processing algorithms, forming a nexus to confront the intricate enigma of glaucoma. As the narrative unfolds, it is evident that within images

* Corresponding author.
E-mail address: robin.augustine@angstrom.uu.se (R. Augustine).

https://doi.org/10.1016/j.pdpdt.2025.104621
Received 4 February 2025; Received in revised form 14 April 2025; Accepted 6 May 2025
Available online 6 June 2025
1572-1000/© 2025 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

lies the potential to decode the "silent thief of sight", ushering in the era where image insights and clinical vigilance converge to preserve visual integrity amidst the intricate interplay of glaucoma's multifaceted image-processing odyssey [13].

Glaucoma remains one of the leading causes of irreversible blindness, with late diagnosis posing a major challenge due to the absence of early symptoms and the limitations of traditional screening methods. Additionally, inconsistencies in optic disc segmentation contribute to variability in clinical diagnosis, affecting the reliability of automated detection systems. Another significant hurdle is the computational inefficiency of existing deep learning models, which limits their practicality for large-scale deployment.

To address these challenges, we propose an approach that enhances segmentation accuracy using U-Net and U-Net++ architectures, improves feature representation through Capsule Networks — effectively preserving spatial hierarchies — and accelerates classification using Extreme Learning Machines (ELM), significantly reducing computational complexity. By integrating these advancements, our model provides a more accurate, efficient, and scalable solution for glaucoma detection, making early screening more accessible and reliable.

The proposed system evaluates median filtering's impact on image quality using multiple metrics. Results indicate 97.88% noise reduction, with noise attraction in only 0.30% of cases. Image degradation was minimal (0.17%). The average noise attenuation of 0.025 signifies substantial improvement, while processing efficiency is reflected in a mean spoiling time of 0.45. Signal-to-Noise Ratio (SNR) peaked at 44.99, underscoring the effectiveness of median filtering in enhancing glaucoma images.

Fundus images were annotated to highlight key glaucoma-related features such as the optic disc and cup. During training, the U-Net model was fine-tuned with comprehensive datasets to recognize subtle glaucomatous traits. Performance assessment showed U-Net achieving a Dice coefficient of 0.8557 and a Jaccard Index of 0.7307, outperforming U-Net+, which recorded values of 0.7670 and 0.6057, respectively. The correlation coefficient, measuring agreement between AI predictions and ground truth, was 0.95 for U-Net and 0.9 for U-Net+, demonstrating strong alignment with clinical observations.

The trained U-Net model facilitates automated segmentation of crucial features such as the optic disc, optic cup, and vascular network. Bland–Altman analysis showed minimal deviation between predicted and actual values, with a mean error of 0.01 mm^2 for U-Net and a standard deviation of 0.02 mm^2 for U-Net+. U-Net consistently achieved an 80% area error below 2.99 mm^2, compared to 4.8 mm^2 for U-Net+. These segmented regions serve as data reservoirs for post-segmentation analysis, revealing essential biomarkers of glaucoma progression.

The study's findings have significant clinical implications, aiding ophthalmologists in glaucoma diagnosis and monitoring. The integration of quantitative assessments [14] with visual insights enhances early detection and evidence-based decision-making, advancing precision in glaucoma management.

1.1. Empowering medical insight: Image processing in healthcare

Strategic treatment planning and enhanced research initiatives. Image processing is the editing and analysis of medical images to extract relevant insights while navigating the complicated environment of medical imaging. Clinicians acquire critical insights into interior anatomical structures and physiological processes [15] using a variety of medical imaging modalities including X-ray, MRI, CT, ultrasound, and PET [16]. Image enhancement techniques such as noise reduction, contrast improvement, and edge sharpening [17] considerably increase image quality, allowing medical experts to comprehend complicated structures and spot probable anomalies more clearly. Segmentation approaches strengthen this process by separating specific regions or components of interest within images [18], making tasks like tumor identification, organ delineation, and blood vessel monitoring more precise. Allowing quantitative analysis of visual data [19]. Texture analysis techniques are important in disease characterization, especially in discriminating between benign and malignant tissue a vital feature in cancer identification [20]. With the incorporation of learning approaches [21], they show promise for automating tasks such as picture classification, segmentation, and even illness prediction, hence improving accuracy in medical image analysis. Registration approaches combine medical images from different images or periods [22], whereas 3D reconstruction converts 2D scans into full 3D visualizations, which aids in surgical planning and teaching [23]. Image fusion combines images from many images, resulting in panoramic representations that improve diagnostic accuracy [24], whereas denoising and restoration techniques remove noise and artifacts from images, increasing diagnostic accuracy [25].

Computer-Aided Detection/Diagnosis (CAD) algorithms help radiologists and clinicians identify potential anomalies in medical images [26], and 3D reconstructions enable virtual exploration and simulation, which fuels surgical planning and practice, fostering a realm of virtual medical simulations [27]. Telemedicine uses image processing to overcome geographical barriers, allowing for remote medical consultations [28], whereas real-time image processing guides interventions like minimally invasive surgeries, providing surgeons with real-time insights that promote precision [29]. Furthermore, image processing contributes to medication development and research by analyzing cellular and molecular images [30].

Glaucoma is a progressive and irreversible optic neuropathy that necessitates early detection and continuous monitoring to prevent vision loss. Diagnostic errors in glaucoma classification — false negatives (FN) and false positives (FP) — carry significant clinical implications. A false negative diagnosis, where a glaucomatous patient is misclassified as normal, can lead to delayed intervention. Without timely treatment, progressive optic nerve damage may continue undetected, increasing the risk of permanent blindness. Additionally, advanced-stage glaucoma, which often presents without symptoms in its early phases, may remain unnoticed until substantial visual field deterioration occurs.

Conversely, false positives, where healthy individuals are incorrectly identified as having glaucoma, contribute to overdiagnosis and unnecessary medical interventions. Such errors can cause undue psychological distress for patients and increase the burden on healthcare systems through unwarranted referrals, additional testing, and excessive treatment regimens. Overuse of intraocular pressure-lowering medications or surgical interventions may expose patients to unnecessary risks and complications.

To mitigate these issues, our study prioritizes model optimization to balance sensitivity and specificity, aiming to minimize FN and FP rates. By incorporating robust feature extraction, deep learning-based segmentation, and extreme machine learning classifiers, our approach seeks to enhance diagnostic accuracy and ensure reliable clinical decision-making. A careful trade-off between sensitivity and specificity is maintained, aligning with the clinical objective of early intervention while reducing unnecessary treatment burdens.

1.2. Navigating the complexities of image segmentation in image processing

To achieve precision and durability in picture segmentation, researchers encounter several technological challenges that highlight the ever-changing nature of image processing [31]. During picture segmentation, a slew of technological obstacles arise, demonstrating the complexities required to generate accurate and dependable findings. These issues reflect the changing landscape of image processing. For starters, segmenting objects with irregular forms or diffuse borders requires sophisticated algorithms that can handle such difficulties [32]. Variations in intensity, lighting conditions, and contrast provide severe obstacles, especially in medical and outdoor scenarios where environmental dynamics impact picture properties [33]. Furthermore, photos from various sources frequently contain noise, artifacts, and distortions,

Table 1
Statement of significance.

Statement of significance	
Significance	Implications
Our research introduces a approach utilizing advanced deep learning techniques, specifically U-Net and U-Net+ architectures, for the detection of glaucoma diseases from retinal images.	This approach offers a more accurate, efficient, and reliable method for identifying glaucoma-related abnormalities, potentially revolutionizing glaucoma diagnosis and improving patient outcomes.
By incorporating feedforward fully connected layers into our model, we achieve superior performance compared to existing approaches, demonstrating the effectiveness of our proposed method in improving early diagnosis and patient outcomes.	Our findings underscore the potential clinical impact of our approach, highlighting its advantages over traditional methods and informing future research directions in the field of medical imaging and informatics.

which can have a substantial impact on the effectiveness of algorithms for segmentation [34].

The scale and size change of objects inside an image canvas, as well as partial occlusion and class imbalance, add to the difficulty of the segmentation process [35–37], Real-time processing needs in dynamic applications such as robots necessitate algorithms capable of rapidly processing images without sacrificing accuracy [38]. The intricacies of anatomical heterogeneity across people add to the difficulty of medical imaging [39]. Interactive segmentation approaches involving user involvement necessitate a careful assessment of beginning circumstances, which presents obstacles, particularly for non-experts [40]. Furthermore, tackling computational complexity and assuring algorithmic efficiency across varied datasets pose new challenges [41]. The absence of exact ground truth annotations for training and assessment, edge leakage, and interclass variability adds to the complex topography of picture segmentation issues [42–44]. Researchers try to solve these challenges and progress in the area of image processing by employing novel methodologies and joint efforts.

1.2.1. Evolving solutions: A dynamic landscape of innovation

Despite these formidable challenges, the landscape of image segmentation in image processing is marked by continual evolution and innovation. Researchers ardently strive to pioneer techniques [45] and methodologies that surmount these hurdles, seeking to elevate the accuracy, reliability, and applicability of segmentation across a spectrum of domains as we delve into the subsequent sections, this narrative of innovation unfolds, shedding light on the strategies and breakthroughs that propel image segmentation toward new horizons [46].

One major challenge is the poor segmentation accuracy of the optic disc and cup, especially in noisy fundus images with varying illumination and contrast. These inaccuracies can significantly affect the reliability of glaucoma diagnosis [47]. Additionally, conventional convolutional neural networks (CNNs) often struggle to preserve spatial relationships crucial for detecting structural changes such as optic disc cupping, a primary indicator of glaucoma progression [48]. Furthermore, many existing models are computationally intensive, making them impractical for deployment in low-resource settings. This underscores the need for lightweight yet accurate architectures that can operate effectively in real time and on limited hardware, enabling broader accessibility and practical implementation in clinical environments.

As depicted in Table 1, the Statement of Significance highlights our research's significant contributions and implications in revolutionizing glaucoma diagnosis and improving patient outcomes.

The rest of the paper is organized as follows: Section 2 discusses Innovations in Glaucoma Detection: Bridging Technology and Clinical Practice, where we explore the integration of advanced methodologies and technologies in glaucoma diagnosis. Section 3 reviews the Related Works, highlighting existing research and deep learning algorithms applied to glaucoma detection. In Section 4, we present the Proposed Scheme, detailing our approach and its architectural components. Section 5 outlines the Experimental Testbench, including datasets and evaluation parameters, while Section 6 provides a thorough Results Analysis of the findings. The implications of these results are discussed in Section 7. Finally, Section 8 concludes the paper, summarizing key contributions and suggesting future research directions in the field of glaucoma detection.

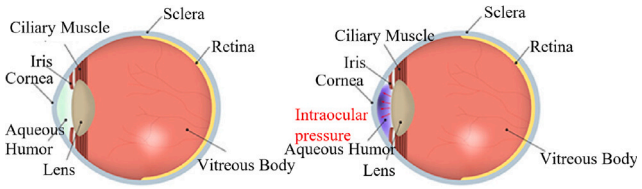

Fig. 1. Comparison between healthy and glaucomatous eyes.

2. Innovations in glaucoma detection: Bridging technology and clinical practice

This study delves into the intricate challenges and innovations at the intersection of image processing, medical diagnosis, and glaucoma detection. As advancements in technology continue to intertwine with medical progress, a synergistic narrative emerges, paving the way for novel methodologies that extend beyond traditional diagnostic paradigms. This ongoing exploration and innovation converge into a coherent narrative that warrants further investigation.

2.1. Glaucoma and retinal image analysis: A pathway to diagnosis

Elevated intraocular pressure within the retinal region serves as a primary precursor to the development of glaucoma. The aqueous humor, a clear fluid responsible for maintaining constant intraocular pressure, can lead to optic nerve damage when its drainage is impaired. This gradual and initially inconspicuous deterioration makes glaucoma the second leading cause of blindness globally, often characterized by a progressive loss of vision. Conventional ophthalmological diagnostic methods, such as tonometry and pachymetry, are essential in identifying the condition. Fig. 1 illustrates the differences between a healthy eye (left) and a glaucomatous eye (right). Imaging techniques, particularly Optical Coherence Tomography (OCT), are crucial in analyzing glaucoma, allowing for monitoring of disease progression and assessment of structural damage. The utilization of advanced imaging modalities is pivotal in glaucoma assessment, providing insights into both structural and functional changes associated with the disease.

2.2. Current diagnostic landscape: Striving for enhanced insight

Current methodologies for glaucoma diagnosis include fluorescein angiography (FA) and Optical Coherence Tomography (OCT), supplemented by clinical examinations and fundus photography [49]. Fluorescein angiograms involve the use of injected fluorescent dye, illuminated under specific wavelengths to produce two-dimensional cross-sectional images of ocular tissues. Visual field analyzers, such as the Humphrey Visual Field Analyzer and Octopus Perimetry, play a key role in detecting functional vision loss, which often complements structural findings. Additionally, direct fundus observation by ophthalmologists, using ophthalmoscopy or slit-lamp biomicroscopy, remains essential for evaluating optic disc changes in a clinical setting.

OCT generates two-dimensional cross-sections using axial scans, allowing for three-dimensional reconstructions of the retinal structure.

Together with fundus photography, these techniques offer a comprehensive diagnostic view. However, despite their efficacy, there remains a pressing need for cost-effective alternatives that facilitate mass screening, early detection, and ongoing disease monitoring [50]. Recent advancements in glaucoma detection through machine learning and deep learning have shown promise but still require improvements in robustness, performance, and simplicity [51].

2.3. Motivation and contribution: Advancing intelligent glaucoma detection

Inspired by these challenges, this research introduces a deep learning model that integrates a Frame Network with visual attention mechanisms for early glaucoma prediction. This innovative approach addresses existing gaps in the current diagnostic landscape by creating a robust glaucoma recognition framework that balances high performance with real-time efficiency. The key contributions of this study include the development of an ensemble model that combines the Frame Network with visual attention layers to accurately predict the onset of glaucoma, distinguishing between healthy and glaucomatous conditions while assessing progression levels.

Furthermore, the inclusion of the U-Net transfer learning network significantly enhances segmentation accuracy, leading to superior predictive performance. By replacing high-complexity deep learning models with visual attention Frame networks, we optimize both robustness and computational efficiency. The shift from traditional fully connected layers to feedforward layers further enhances classification capabilities. Rigorous empirical validation demonstrates the superiority of the proposed algorithm over existing methods, confirming its efficacy.

Through the integration of advanced deep learning architectures and optimization of segmentation efficiency, this research charts a transformative path for early glaucoma prediction. Its multifaceted contributions drive glaucoma detection toward greater accuracy, efficiency, and reliability, revitalizing the field of ocular health prognosis [20].

As technological advancements continue to parallel developments in medicine, a new horizon emerges that fuses deep learning capabilities with the precision of medical image processing [52]. This study presents a paradigm through the innovative sandwiched deep learning model, which combines Frame Network and visual attention layers [53]. This fusion transcends traditional glaucoma prediction boundaries, opening pathways for early intervention and enhanced diagnostic insights. The model's predictive power extends across various stages of glaucoma, supported by the precise segmentation capabilities of the U-Net transfer learning network. Computational efficiency, backed by visual attention Frame networks, aligns with enhanced classification through feedforward layers. Overall, this research narrates a compelling story of empirical validation, showcasing its advantages over contemporary approaches.

The journey embarked upon, driven by innovation and grounded in pressing medical needs, weaves together the narratives of image processing, medical diagnostics [54], and glaucoma detection. This voyage acknowledges the intricate challenges facing the field while heralding a new era that unites computational power with clinical insight. As subsequent sections delve deeper into methodologies, experimental designs, and results, the harmony of science and technology resonates, ushering in a transformative period for early glaucoma detection, enhanced diagnostic precision, and a renewed commitment to ocular health.

3. Related works

The pursuit of precision in Glaucoma detection has ignited a surge of research endeavors, culminating in a spectrum of sophisticated methodologies that synergize the potency of deep learning and cutting-edge medical imaging techniques. In this section, we traverse a landscape of notable contributions that unveil the trajectory toward early Glaucoma identification and delineate the latest advancements.

Manassakorn et al. [55] forged a path by devising a Glaucoma diagnostic network predicated on optical coherence tomography angiography imaging. This pioneering approach hinged on quantitative parameters, culminating in clinical diagnostic prowess with a commendable balance between sensitivity (88.9%) and specificity (89.6%). Notably, this framework orchestrates an accelerated training phase while maintaining computational complexity. Liao et al. [56] ventured into the realm of clinically interpretable Convolutional Neural Networks (ConvNets) for robust Glaucoma diagnosis. The resulting Multi-Layer Average Pooling architecture (M-LAP) propounds pooling insights across multiple scales, thus accentuating diagnostic accuracy (AUC 0.88). Challenges in conveying high-resolution feature maps underscore the ever-evolving nature of this endeavor. Li et al. [57] ushered in the era of attention-driven ConvNets for Glaucoma diagnosis. Operating in a semi-supervised fashion, AG-CNN artfully merged attention prediction, pathological region localization, and Glaucoma classification subnets. This framework emerged victorious in accuracy, sensitivity, specificity, AUC, and F1-score while concurrently mitigating computational complexity.

George et al. [58] navigated the three-dimensional domain, devising an attention-guided 3D Deep Learning model for Glaucoma detection. Fueled by the profundity of 3D Optical Coherence Tomography (OCT) cubes, their model, empowered by grad-CAM attention maps, unfurled enhanced resilience and performance. The formidable challenge of real-time deployment, however, remains an ongoing pursuit. Zhao et al. [59] introduced MFPPNet, a multi-scale convolutional network engineered for unsupervised fundus image representation learning. The resultant glaucoma classification framework underscored superior MAE and AUC outcomes, even as it grapples with the nuanced intricacies of heightened computational complexity. Islam et al. [60] orchestrated an amalgamation of deep learning techniques, embracing EfficientNet, MobileNet, DenseNet, and GoogLeNet, for automated Glaucoma classification. This ensemble emerged triumphant in augmenting test accuracy, F1-score, and ROC AUC. Efficiency in resource utilization emerged as a hallmark attribute.

Shoukat et al. [61] harnessed CNN architectures for early Glaucoma identification, enlisting a series of filters to elevate fundus image quality for classification. Through the potent utilization of the EfficientNet architecture, their model achieved state-of-the-art accuracy, sensitivity, and specificity on the DRISHTI-GS dataset. This accomplishment, however, yielded elevated computing overhead. Parashar and Agrawal [62] embarked on a distinctive trajectory, unraveling a 2D-DWT-based automated categorization technique for Glaucoma. The extraction of distinctive attributes from DWT-deconstructed components propelled image classification accuracy to an impressive 91.67%. Yet, the simplicity and accuracy of this approach are counterbalanced by its limited scalability to larger datasets.

Ovreiu et al. [63] plunged into densely linked neural networks for early Glaucoma detection, achieving a noteworthy accuracy of 95.6% and F1-score of 0.97%. While buoying the prospects of population screening, the performance of this approach appeared sensitive to dataset expansion. Suttapakti et al. [64] introduced Multi-directional Texture Feature Extraction (MTFE), augmenting Glaucoma classification accuracy through meticulous texture information extraction. With a categorization accuracy of 90.44%, this framework's forte lies in expedited training intervals. Nonetheless, reduced-resolution images and the augmentation of geometric attributes unveil layers of nuanced complexity [65].

In the domain of glaucoma detection methodologies, an intricate comparative analysis was conducted to assess the effectiveness of various approaches, as succinctly outlined in Table 2. Each method underwent evaluation based on distinct performance metrics, revealing both advantages and limitations. Manassakorn et al. [55] "GlauNet" method demonstrated a sensitivity of 88.9% and specificity of 89.6%, showcasing proficiency in glaucoma identification. W. Liao et al.'s [56] "Convolutional Networks (ConvNet)" achieved an AUC of 0.88, offering

Table 2
Comparison of glaucoma detection methodologies.

Ref	Proposed methodology	Results achieved
[55]	GlauNet	Sensitivity (88.9%), Specificity (89.6%)
[56]	Convolutional Networks (ConvNet)	AUC (0.88)
[57]	Attention-Guided CNN (AG-CNN)	Accuracy (96.2%), Sensitivity (95.4%), Specificity (96.7%), AUC (0.983), F1-Score (0.954)
[58]	End-to-End 3D DL with Attention	RMSE (13.403), MAE (7.954), MDAE (3.615)
[66]	Multi-Fusion Pre-Processing Net	MAE (0.0606), AUC (0.88)
[60]	Multi-Model Deep Learning	Accuracy (0.9652), F1-Score (0.9512), ROC AUC (0.9574)
[61]	Convolutional Neural Networks	Accuracy (98.0%), Sensitivity (95.19%), Specificity (94.0%)
[62]	2D Discrete Wavelet Transform	Accuracy (91.67%)
[63]	Densely Connected Neural Nets	Accuracy (95.6%), F1-Score (0.97)
[64]	Multi-Task Feature Extraction	Accuracy (90.44%)
[68]	U-Net models with CNN architectures (Inception-v3, VGG19, and ResNet50)	highest accuracy of 99.53% in segmenting OD using the RIM-ONE dataset
[69]	CNN architectures (Inception-v3, VGG19, and ResNet50)	highest accuracy of 98.52% for the ACRIMA fundus image dataset

a robust foundation for precise glaucoma diagnosis. L. Li et al.'s [57] "Attention-Guided CNN (AG-CNN)" yielded accurate diagnostics with 96.2% accuracy, 95.4% sensitivity, 96.7% specificity, and a noteworthy AUC of 0.983, highlighting adeptness across diverse cases. Y. George et al.'s "End-to-End 3D DL with Attention" exhibited high robustness, evidenced by RMSE of 13.403, MAE of 7.954, and MDAE of 3.615, showcasing accuracy across predictions. P. Kunumpol et al.'s [66] "Multi-Fusion Pre-Processing Net" achieved MAE of 0.0606 and AUC of 0.88, exemplifying accuracy and effective prediction foundations. M. T. Islam et al.'s [60] "Multi-Model Deep Learning" presented a balanced performance, with 0.9652 accuracy, 0.9512 F1-score, and 0.9574 ROC AUC, achieving a trade-off. A. Shoukat et al.'s [61] "Convolutional Neural Networks" method demonstrated promise with 98% accuracy, 95.19% sensitivity, and 94% specificity, signifying potential accuracy. D. Parashar et al.'s "2D Discrete Wavelet Transform" achieved 91.67% accuracy, balancing accuracy and complexity. S. Ovreiu et al.'s "Densely Connected Neural Nets" yielded 95.6% accuracy and 0.97 F1-score, portraying cost-effective performance. U. Suttapakti et al.'s "Multi-Task Feature Extraction" showcased 90.44% accuracy with reduced training time, though limitations existed for lower-resolution images. This exhaustive analysis imparts insights into strengths and limitations, aiding decision-making when selecting glaucoma detection methods based on requirements and trade-offs [67].

The corpus of seminal works traversed above serves as an eloquent testimony to the dynamic tapestry of Glaucoma detection. Through an intricate dance of innovative architectures, attention mechanisms, and fusion methodologies, the realm of early Glaucoma identification stands reinvigorated [70]. As we navigate these advancements, a resounding narrative emerges one that underscores our unwavering commitment to eradicating Glaucoma's insidious impact while propelling us closer to a future where timely diagnosis and intervention wield unparalleled efficacy.

Several confounding factors can impact the accuracy and reliability of AI-driven glaucoma diagnosis. Variations in imaging quality, including differences in optical coherence tomography (OCT) device specifications, image resolution, lighting conditions, and contrast variations, can introduce inconsistencies in AI-based predictions. These discrepancies must be considered when training and validating models across diverse clinical settings.

Additionally, population-based variability in optic disc morphology, such as differences in optic disc size and cup-to-disc ratio across ethnic groups, may lead to inherent biases in AI-generated predictions. This highlights the need for diverse, multiethnic datasets to improve model generalizability and fairness.

Co-existing ocular pathologies such as high myopia, diabetic retinopathy, and optic neuropathies can present structural changes that mimic glaucomatous damage, increasing the likelihood of false positives. Moreover, age-related physiological changes, particularly the natural thinning of the retinal nerve fiber layer (RNFL), complicate the differentiation between normal aging and early glaucoma, posing a significant challenge in clinical decision-making.

To mitigate these confounding factors, we emphasize the importance of cross-validation using independent datasets from varied clinical environments. Furthermore, model interpretability techniques should be integrated to allow clinicians to assess AI predictions in conjunction with traditional diagnostic methods. The tables now include footnotes specifying dataset limitations, and figures have been enhanced to ensure clarity and better visualization of key findings.

Visual attention mechanisms significantly improve feature localization and segmentation accuracy by focusing on the most relevant regions of an image. However, this enhancement comes at a cost—higher computational complexity, increased memory usage, and longer inference times. Compared to standard CNN architectures, attention-based models require additional processing to compute attention maps, making them less suitable for real-time applications in resource-constrained environments.

To mitigate these challenges, several optimization strategies can be employed. Model pruning and quantization techniques help reduce the number of parameters and lower memory footprint without significantly compromising accuracy. Knowledge distillation, where a compact model is trained using the outputs of a larger model, can retain performance while improving efficiency. Additionally, lightweight attention mechanisms, such as self-attention modules with lower computational loads, may offer a compromise between performance and inference speed.

Thus, while visual attention frame networks provide clear advantages in segmentation quality and diagnostic reliability, their deployment in large-scale glaucoma screening programs requires careful consideration of computational efficiency. Future work should explore adaptive attention models that dynamically adjust complexity based on available resources, ensuring a balance between accuracy and feasibility in clinical practice.

4. Methods

In this section, we delve into the intricate materials and methodologies underpinning our research endeavors. Our approach is grounded in the careful selection of databases that cater to the unique demands of optic nerve segmentation and retinal imaging analysis. We highlight the rationale behind our choices and discuss the advantages of these databases over others. Subsequently, we outline the datasets adopted for evaluating our proposed algorithm, shedding light on their distinctive attributes and significance. Moreover, we elucidate the preprocessing techniques applied to enhance image clarity and standardize input dimensions. This section offers a comprehensive understanding of the foundational pillars supporting our investigative journey.

4.1. International DRIONS-DB (optic nerve segmentation database)

As [71,72] have noted, the DRIONS-DB database was chosen for optic nerve segmentation research due to its specialized design to satisfy the specific requirements of this discipline. This resource contains precisely annotated retinal images with manually drawn optic nerve head borders, resulting in a rich dataset for the development and testing of cutting-edge segmentation algorithms. A significant aspect of DRIONS-DB is the availability of thoroughly managed ground truth annotations. This internal validation process promotes the development of precise and resilient approaches for optic nerve segmentation. Furthermore, DRIONS-DB is a crucial tool for comparison analysis, allowing researchers to examine alternative segmentation algorithms and approaches side by side, hence expediting development in optic nerve analysis and broader medical applications.

4.2. High-resolution fundus (HRF) database

The HRF database's inclusion in research endeavors is motivated by its unique characteristic of offering high-resolution fundus images collected using modern fundus cameras, assuring the availability of high-quality images required for extensive retinal structure analysis. These high-quality images are critical in detecting various retinal illnesses because they allow researchers and physicians to precisely examine complex features. Furthermore, the HRF database's popularity stems from its wide range of retinal diseases, which includes disorders such as diabetic retinopathy and glaucoma. This variety enables both researchers and clinicians to participate in the early identification and ongoing monitoring of various disorders, expanding the field of research and therapeutic applications [73].

The HRF database is inclusive, containing photos from a diverse range of individuals with varied disease severity levels and stages. This heterogeneous dataset reflects real-world complexity, giving researchers significant insights into multiple clinical environments. Aside from research, the HRF database is useful in clinical settings since it contains a large number of retinal images [74]. This comprehensive resource acts as a catalyst for the development of automated illness detection and categorization systems, resulting in improved diagnosis accuracy and patient care. While the specialized scope, exact annotations, and different datasets impact the selection of databases such as HRF, researchers and practitioners should also consider their individual study aims and application needs. For the most current data, review recent literature and appropriate sites.

4.3. Dataset selection and pre-processing for proposed algorithm evaluation

To systematically verify the efficacy of the proposed approach, three publically accessible datasets were used: DRISHTI-GS, the Optic Nerve Segmentation Database (DRIONS-DB), and the High-Resolution Fundus (HRF). The original image resolutions vary across datasets, with DRISHTI-GS containing images at 2048 × 2048 pixels, DRIONS-DB at 600 × 400 pixels, and HRF providing high-resolution images of 3504 × 2336 pixels. To ensure uniformity and optimize computational efficiency, all images were resized to a standard dimension of 256 × 256 pixels before being processed by the model. This resizing step preserves essential anatomical structures, particularly the optic disc and cup, which are crucial for glaucoma detection. Maintaining a consistent input size enhances the model's ability to learn robust feature representations while ensuring computational scalability across different datasets.

Visualization of Dataset and Model Input: A visual representation of the datasets and model inputs is illustrated in Fig. 2, offering a glimpse into the diverse range of images harnessed for evaluation.

Table 3 provides a comprehensive overview of the datasets employed in this study, detailing both the original and post-augmentation distributions of normal and glaucomatous samples. These publicly

Fig. 2. Sample images from the DRISHTI-GS, DRIONS-DB, and HRF datasets used for training and testing the proposed U-Net and U-Net+ models for optic disc segmentation and glaucoma classification.

available retinal image databases — DRISHTI-GS, DRIONS-DB, and HRF — serve as the foundational data sources for training and evaluating the proposed deep learning framework for automated glaucoma detection.

Due to the relatively limited number of annotated clinical samples in each dataset, particularly in the case of the HRF dataset, data augmentation techniques were employed to artificially expand the training set. Augmentation strategies included affine transformations (rotation and flipping), contrast normalization, and Gaussian noise injection. These transformations aim to simulate variations commonly encountered in clinical imaging, thereby improving the model's generalization and robustness to noise, orientation, and brightness variability.

While augmentation significantly boosts performance by increasing the diversity of the training set, it is important to acknowledge that such transformations do not introduce new pathological cases or unseen phenotypes. Instead, they enhance the model's ability to learn invariant features and prevent overfitting. Therefore, although the number of training samples increases substantially post-augmentation, the clinical diversity of the dataset remains unchanged.

Our approach carefully balances augmentation with validation on non-augmented, real-world test samples to ensure reliable evaluation. Furthermore, we emphasize that augmentation complements but does not substitute the need for large-scale, heterogeneous, real patient datasets that better reflect the range of glaucoma manifestations across populations and imaging conditions.

4.4. Functional requirements

The proposed glaucoma detection model was implemented in Python, utilizing essential deep-learning libraries and tools to facilitate rapid model development and training. The primary framework employed was TensorFlow 2.10, which, in conjunction with Keras 5.5.8, provided a high-level API that simplifies the creation and training of models, particularly for the U-Net architecture and Extreme Learning Machines (ELM). For image preprocessing and augmentation tasks, OpenCV 1.10 was utilized to enhance image quality and diversity. Additionally, the integration of Capsule Networks (CapsNet) with U-Net significantly improved the model's ability to capture complex features in the images, thereby enhancing diagnostic accuracy.

The implementation was executed on a high-performance PC workstation equipped with an Intel I9 CPU and 16 GB of RAM. This configuration efficiently managed the initial stages of preprocessing, dataset handling, and other computational tasks that did not necessitate extensive parallel processing. For deep learning tasks, particularly model training and optimization, an NVIDIA Titan GPU with 256 GB of RAM was employed. The utilization of GPUs is crucial for accelerating the training of neural networks [75], as they are designed to handle large-scale parallel computations required by frameworks like TensorFlow and Keras. In this setup, the NVIDIA Titan GPU was primarily responsible for the intensive computational tasks associated with model training, while the Intel I9 CPU supported auxiliary operations, such as preprocessing and dataset management. Consequently, both the CPU and GPU played significant roles in the implementation, with the GPU serving as the main component for deep learning computations.

Table 3
Summary of public fundus image datasets before and after augmentation.

Sl. No.	Dataset	Original normal	Original glaucoma	Augmented normal	Augmented glaucoma
1	DRISHTI-GS	75	26	380	122
2	DRIONS-DB	63	47	390	220
3	HRF	10	20	198	180

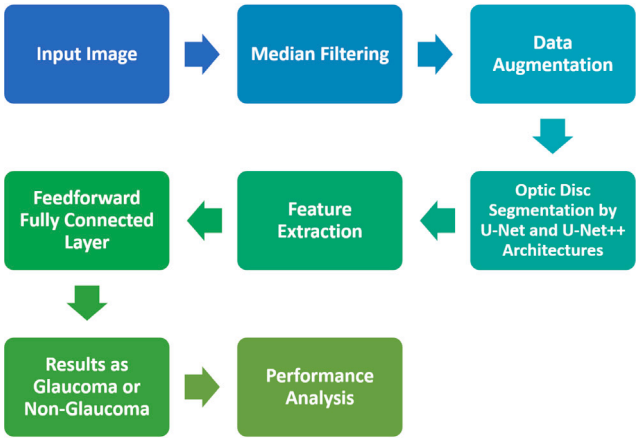

Fig. 3. Proposed architecture for glaucoma detection.

5. Proposed advancements in glaucoma detection

This section delves into the intricacies of the operational blueprint underpinning innovative ensemble architecture. The holistic architecture, outlined in Fig. 3, represents a strategic amalgamation of five seamlessly orchestrated stages: data pre-processing, data augmentation, segmentation, feature extraction, and feedforward layers. Together, these stages synergize to propel the efficacy of our proposed system to unprecedented heights.

The route to efficient glaucoma diagnosis begins with thorough data pre-processing and augmentation of input images, which serve as the foundation for further analysis. This fundamental step is critical because it improves training performance by improving data quality, standardizing formats, and correcting anomalies. Augmentation techniques add to the dataset, ensuring strong model generalization and robustness to perturbations. Following this first step, the process shifts to the critical component of optic disc segmentation. Using the power of cutting-edge U-NET networks, this stage effortlessly separates the optic disc, unraveling its anatomical complexities and enabling accurate localization. With optic disc fragmentation as the fulcrum, the design effortlessly moves to feature extraction and classification. This step, orchestrated by attention-driven Frame networks, carefully isolates relevant characteristics that capture diagnostic subtleties. This combined system, enhanced by extreme learning machines, achieves unprecedented diagnostic accuracy and rapid decision-making. The pinnacle of extracted characteristics occurs in the domain of feedforward layers, where synthesized information is meticulously classified. Here, detailed patterns are identified, distinguishing between normal and glaucoma-affected cases. After classification, the findings are thoroughly analyzed, with measures such as sensitivity, specificity, and accuracy, used to assess system performance. This investigation sheds light on the suggested methodology's efficacy and reliability for identifying glaucoma [76]. Clinicians are enabled to make educated decisions about patient diagnoses and treatment regimens, which improves patient care and contributes to better visual health outcomes. To ensure continuous progress, the system goes through iterative optimization cycles, enhancing its ability to perform based on clinician feedback and ongoing examination of system outputs. Rigorous validation and verification processes are used to assure the dependability and authenticity of system outputs, establishing trust in the system's capabilities. Following adequate testing and verification, the system is implemented for clinical usage, simplifying diagnostic processes and allowing for early diagnosis and quick intervention in glaucoma patients. Finally, the broad deployment of this cutting-edge technology has a tremendous influence on patient treatment, ushering in a new age of early identification, intervention, and better results in glaucoma.

- Data Pre-processing and Augmentation: The journey commences with the meticulous pre-processing and augmentation of input images. This critical phase optimizes the training performance by enhancing data quality, standardizing formats, and rectifying anomalies. Augmentation techniques further enrich the dataset, ensuring robust model generalization and resilience to variations.
- Optic Disc Segmentation: A pivotal cornerstone of our proposed system, the optic disc segmentation stage is executed with the potency of state-of-the-art U-NET networks. Harnessing the power of revolutionary U-NET architecture, this stage seamlessly segregates the optic disc, unraveling its anatomical intricacies and fostering precise localization.
- Feature Extraction and Classification: With optic disc segmentation as a fulcrum, the architecture seamlessly transitions into the domain of feature extraction and classification. Attention-driven Frame networks orchestrate this phase, intricately designed to extract pertinent features encapsulating diagnostic nuances [77]. Augmented by extreme learning machines, this amalgamated system engenders unparalleled diagnostic accuracy and expeditious decision-making.
- Feedforward Layers: The culmination of extracted features transpires within the realm of feedforward layers. Here, the synthesized information is subjected to meticulous classification, deftly discerning intricate patterns and discriminating between normal and Glaucoma-affected instances.

Algorithm 1 Preprocessing Pipeline for Fundus Images to Aid Glaucoma Detection via Cup-to-Disc Ratio Analysis

Require: Raw fundus image I_{raw}
Ensure: Preprocessed image I_{prep} suitable for optic cup and disc segmentation
1: **Red Channel Extraction:** $I_r \leftarrow \text{ExtractRedChannel}(I_{raw})$
2: **Background Estimation:** $I_{bg} \leftarrow \text{MedianFilter}(I_r, \text{kernel}=3\times 3)$
3: **CLAHE Enhancement:** $I_{clahe} \leftarrow \text{ApplyCLAHE}(I_r)$
4: **Gaussian Smoothing:** $I_{smooth} \leftarrow \text{GaussianFilter}(I_{clahe}, \text{kernel}=5\times 5, \sigma = 1.5)$
5: **Shade Correction:** $I_{shade} \leftarrow \text{ShadeCorrection}(I_{smooth}, I_{bg})$
6: **Global Normalization:** $I_{norm} \leftarrow \frac{I_{shade} - \mu(I_{shade})}{\sigma(I_{shade})}$
7: **Intensity Scaling:** $I_{prep} \leftarrow \frac{I_{norm} - \min(I_{norm})}{\max(I_{norm}) - \min(I_{norm})}$
8: **return** I_{prep}

The preprocessing pipeline presented in Algorithm 1 is designed to prepare raw fundus images for high-accuracy segmentation of the optic disc and optic cup, which are critical for glaucoma diagnosis via cup-to-disc ratio (CDR) analysis. Each step in this pipeline plays a crucial role in enhancing image quality, reducing noise, and normalizing visual features across the dataset, thereby improving the robustness of deep learning models trained on these images.

The algorithm begins with red channel extraction from the original RGB fundus image. This is based on the clinical observation that the red channel offers higher contrast for the optic disc and cup boundaries due to its sensitivity to vascular and retinal structures. By isolating the red channel, irrelevant background information is minimized, and key features become more prominent for subsequent processing stages.

Following this, median filtering with a 3×3 kernel is applied to estimate the background. Median filtering is particularly effective in removing salt-and-pepper noise and in smoothing local intensity variations while preserving anatomical edges, making it suitable for retinal imaging.

Next, the image undergoes contrast enhancement through Contrast Limited Adaptive Histogram Equalization (CLAHE). This technique improves the visibility of optic nerve head features, particularly in poorly illuminated or overexposed areas, by redistributing pixel intensities locally within image tiles, while limiting noise amplification.

Gaussian filtering is then employed with a 5×5 kernel and standard deviation $\sigma = 1.5$ to further reduce high-frequency noise. This smoothing operation targets sensor-related and lighting-induced noise, providing a cleaner signal for edge and structure detection without degrading important retinal features.

Shade correction is applied to mitigate uneven illumination across the retinal field. This is done by subtracting the estimated background from the smoothed image, effectively normalizing the brightness of the foreground structures and preserving their local contrast.

To standardize image intensity, global contrast normalization is performed by subtracting the mean and dividing by the standard deviation of pixel intensities. This step ensures consistent image representation across samples and reduces the variability due to lighting conditions or patient-specific retinal pigmentation.

Finally, the pixel intensities are scaled to a fixed range of $[0, 1]$ through min–max normalization. This is essential for training stability and compatibility with deep learning architectures that are sensitive to input scale variations.

This structured pipeline ensures the delivery of high-quality, denoised, and contrast-enhanced fundus images suitable for automated segmentation. During the experimental setup, the dataset is split into 80% for training and 20% for testing, with 10% of the training data set aside for validation. This approach balances generalization and overfitting control, enhancing the overall performance of glaucoma detection models.

While U-Net and U-Net++ are well-established segmentation models, the novelty of our approach lies in their enhanced integration with Capsule Networks and Extreme Learning Machines (ELM) to create a more robust and efficient glaucoma detection framework. Unlike conventional CNNs, Capsule Networks capture complex spatial relationships within the optic disc, improving feature extraction and segmentation accuracy. Additionally, ELM-based classification provides a significant advantage over deep networks by enabling faster convergence, reducing computational costs, and improving real-time applicability without compromising diagnostic performance. These enhancements ensure a more scalable and clinically viable solution for glaucoma detection.

These meticulously orchestrated stages foster a seamless symphony of operations. The resulting ensemble architecture epitomizes the synergistic marriage between deep learning prowess and innovative methodologies. It unveils a transformative paradigm in Glaucoma detection, poised to catalyze early identification and intervention, ultimately fostering enhanced patient care and visual health.

5.1. Image augmentation

After scaling the photographs to a standardized dimension, the next hurdle is gathering a large and varied dataset to properly train deep neural networks [78]. To solve this difficulty, the use of data augmentation techniques becomes critical, constituting a key approach in current machine learning. Data augmentation refers to a series of changes aimed at extending the dataset without introducing bias or altering the original images. These modifications, which include horizontal and vertical flipping, rotation, translation, and other operations, add up to an enhanced dataset that captures a wide range of real-world variances.

In our study, affine transformations are used as a geometric modification approach to retain parallel lines and ratios, guaranteeing that the modified images are anatomically correct. The use of affine transformations considerably enhances the dataset's size and variety, which is critical for avoiding overfitting and improving the model's generalization capabilities. This augmentation procedure gradually incorporates variation into the dataset, giving the neural network a broader set of stimuli to learn from. By developing numerous copies of each image via controlled changes, the model grows more adaptable and capable of managing differences that may arise in real-world situations. The augmentation procedure, which employs affine transformations, has proven beneficial in this study by supplementing the current dataset with a large number of extra photos.

Upon resizing the images to a standardized dimension, the subsequent challenge emerges in acquiring a substantial and diverse dataset for the effective training of deep neural networks. This necessitates the utilization of data augmentation techniques, a pivotal strategy in modern machine learning. Data augmentation encompasses a spectrum of transformations designed to expand the dataset without introducing any bias or distorting the original images [79]. These transformations include operations such as horizontal and vertical flipping, rotation, translation, and more, each contributing to an enriched dataset that encapsulates various real-world variations.

The DRISHTI-GS dataset initially contained 101 images (26 glaucoma, 75 normal), DRIONS-DB had 110 images (47 glaucoma, 63 normal), and HRF included 30 images (20 glaucoma, 10 normal). Augmentation significantly increased the dataset size, leading to a more balanced distribution: DRISHTI-GS (380 normal, 122 glaucoma), DRIONS-DB (280 normal, 220 glaucoma), and HRF (198 normal, 180 glaucoma). This balanced dataset reduces bias and enhances model robustness.

We employed median filtering as a primary technique for noise reduction, achieving a noise reduction efficiency of 97.88%. This method was chosen due to its ability to effectively reduce salt-and-pepper noise while preserving edges, which is crucial for maintaining important structural details in medical images. The efficacy of median filtering was validated through a peak signal-to-noise ratio (SNR) of 44.99 and a mean attenuation value of 0.025, demonstrating significant improvements in image quality with minimal distortion [80]. The rationale for selecting median filtering is particularly relevant in medical imaging, where preserving anatomical integrity is essential for accurate analysis and diagnosis.

In addition to median filtering, we standardized the images by resizing them to ensure uniformity across the dataset. This resizing process is critical for maintaining a consistent input shape, which is necessary for the effective training of convolutional neural networks [77]. Furthermore, we applied standardization to normalize pixel values, enhancing the model's training efficiency and overall performance in the segmentation tasks.

5.1.1. U-Net and U-Net+ architectures

The U-Net architecture is specifically designed for biomedical image segmentation. It consists of a contracting path, where feature maps are captured through convolutional layers, followed by a symmetric expanding path that enables precise localization. The architecture's hallmark is its skip connections, which link corresponding layers in the contracting and expanding paths, allowing the model to retain spatial information while capturing contextual features. This structure is particularly effective for tasks requiring high accuracy in pixel-level segmentation, such as identifying glaucomatous changes in retinal images.

Building on the U-Net, the U-Net+ architecture introduces several enhancements to improve segmentation performance. It incorporates nested skip pathways and attention mechanisms, which facilitate better feature reuse and more refined localization of relevant features. These modifications allow the U-Net+ architecture to effectively leverage multi-scale features, improving its ability to detect subtle changes in retinal images that are critical for early glaucoma detection. In our proposed approach, we utilize the U-Net and U-Net+ architectures to accurately segment retinal images and identify features associated with glaucoma. By employing these architectures, we aim to enhance the model's ability to detect glaucomatous changes, thereby providing a robust tool for clinicians in diagnosing and monitoring this condition.

To facilitate accurate and automated optic disc segmentation, we leverage the U-Net architecture, a convolutional neural network tailored for image segmentation tasks. This architecture features a distinctive arrangement comprising two fundamental components: the contracting path and the expansive path.

The contracting path encompasses a series of successive layers, each comprising a 3 × 3 convolutional layer followed by a rectified linear unit (ReLU) activation function and a pooling layer. This architecture is reminiscent of conventional convolutional neural networks and is repeated iteratively.

Accurate segmentation of the optic disc and cup is crucial for glaucoma diagnosis, particularly in determining the cup-to-disc ratio (CDR), a critical clinical parameter. U-Net's convolutional operations inherently lead to a loss of border pixels, often mitigated through padding strategies such as reflection padding. This technique preserves edge continuity by extending natural image features, minimizing artificial discontinuities that may arise from zero-padding or constant padding. Border inaccuracies can introduce measurement variability in CDR calculations; however, in clinical practice, minor edge distortions typically do not significantly impact glaucoma diagnosis when the central optic nerve head is correctly segmented.

To further reduce this issue, our implementation integrates skip connections, which retain spatial information and enhance border refinement. The use of post-processing techniques, including morphological operations and active contour models, further refines segmentation outputs. Despite potential minor losses at the periphery, our Dice Similarity Coefficient (DSC) and Jaccard Index values demonstrate strong agreement between predicted and ground truth segmentations, ensuring that CDR calculations remain within clinically acceptable diagnostic thresholds. Future work may explore more advanced strategies, such as hybrid deep learning models incorporating transformer-based architectures, to further enhance boundary preservation and segmentation precision

One of the critical anatomical indicators for glaucoma progression is the *neuroretinal rim area*, which is computed as the difference between the optic disc area and the optic cup area:

$$\text{Rim Area} = \text{Disc Area} - \text{Cup Area} \quad (1)$$

A reduced rim area is indicative of rim thinning, a hallmark feature of glaucomatous optic neuropathy. Accurate quantification of this parameter is essential for effective glaucoma detection and staging, as it provides additional insight beyond the commonly used cup-to-disc ratio (CDR).

To improve the precision of optic disc and cup segmentation, the red channel of the RGB fundus image is selectively utilized. The optic disc typically exhibits a higher intensity response in the red channel due to its rich vascular and neural structure. This channel enhances the contrast between the disc and surrounding retina, facilitating better boundary localization and aiding in the detection of rim pallor—an early sign of glaucomatous damage.

Incorporating red channel intensity into the preprocessing and segmentation stages allows for more accurate delineation of disc margins, leading to improved estimation of both the rim area and CDR. This enhancement not only refines structural feature extraction but also supports more robust classification performance in automated glaucoma detection models.

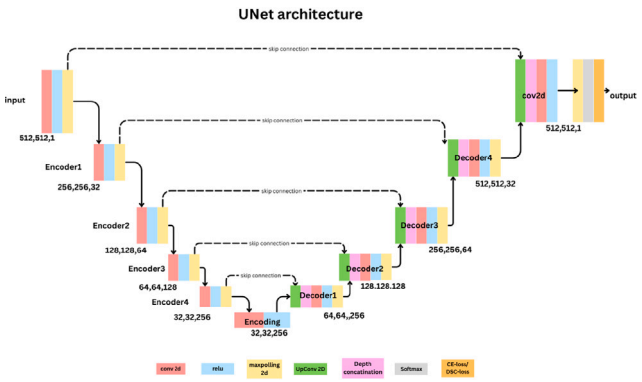

Fig. 4. U-Net architecture for glaucoma segmentation [81].

Fig. 4 illustrates the intricate U-Net architecture employed for optic disc segmentation. This architecture showcases a unique design featuring both a contracting path and an expansive path. The contracting path encompasses multiple successive layers, each consisting of a 3 × 3 convolutional layer followed by a ReLU activation function and a pooling layer. This arrangement captures and abstracts essential image features. The expansive path, on the other hand, employs up-convolutional layers to upsample the feature maps, which are then concatenated with corresponding feature maps from the contracting path. The concatenated information is refined through successive convolutional layers, resulting in the generation of finely segmented optic disc regions. This architecture, resembling a U-shape, facilitates contextual information propagation, enabling the network to make informed segmentation decisions based on a comprehensive understanding of the image context. The careful orchestration of these pathways ensures the U-Net's efficacy in accurately segmenting the optic disc, a crucial step toward effective glaucoma diagnosis.

Fig. 4 illustrates that the U-Net design takes a methodical approach to picture segmentation, using an encoder–decoder network structure supplemented with skip connections. This design carefully incorporates two crucial components: the encoder, which down-samples the input picture to extract high-level features, and the decoder, which inversely up-samples these features to produce precise pixel-by-pixel segmentation masks. Crucially, skip connections build direct interconnections between the encoder and decoder's equivalent layers, allowing for the merging of local and global contextual data.

U-Net++ enhances the traditional U-Net architecture by introducing dense skip connections that facilitate progressive feature refinement across multiple convolutional layers. These nested skip pathways promote better feature fusion, improving the network's ability to extract both fine-grained details and high-level structural information, particularly in complex medical imaging tasks such as optic disc segmentation.

One key advantage of U-Net++ is its ability to mitigate information loss during downsampling and upsampling operations, ensuring a more stable flow of multi-scale contextual information. This is particularly beneficial when segmenting the optic nerve head, where precise boundary delineation is crucial for clinical assessment. Our experimental results indicate that U-Net++ achieves higher segmentation accuracy in cases with substantial image noise and anatomical variability.

However, the architectural complexity of U-Net++ introduces increased computational overhead due to its expanded number of layers and additional feature refinement steps. In scenarios where the base U-Net already performs well — such as in our dataset where clear optic disc boundaries are present — the incremental improvements offered by U-Net++ may not justify the additional computational cost.

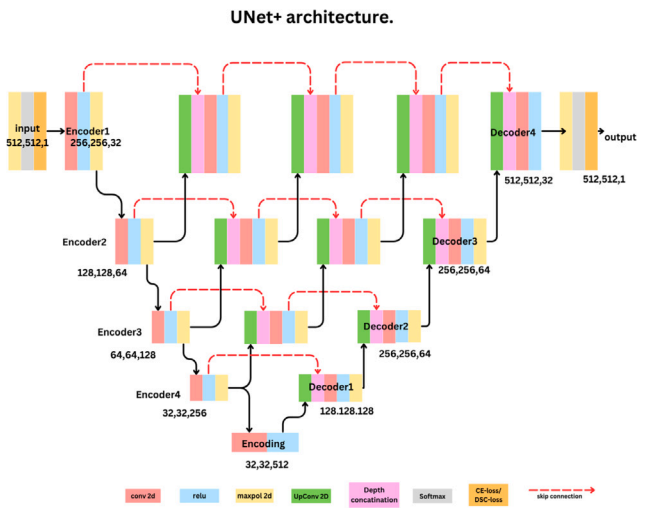

Fig. 5. U-Net+ architecture for glaucoma segmentation [82].

Therefore, while U-Net++ provides tangible benefits in challenging segmentation scenarios, its implementation should be carefully evaluated against the available computational resources and the quality of input images. Our study highlights the effectiveness of both U-Net and U-Net++, demonstrating that the choice between them should be guided by dataset characteristics and real-world deployment constraints.

Fig. 5 illustrates the U-Net+ improves on this basis by including dense connections in both the encoder and decoder blocks. This invention promotes substantial interlayer communication, allowing for smooth feature sharing at all network levels. By tackling issues like the vanishing gradient problem, U-Net+ improves gradient propagation during training, resulting in more robust and efficient learning. the U-Net design matures into U-Net+ by including dense connections, resulting in a complete network structure that optimizes information flow and encourages collaborative learning among network levels.

U-Net and U-Net++ serve as effective segmentation architectures, their structural differences play a crucial role in our approach. U-Net employs a single skip connection per layer, which facilitates information flow but may lead to the loss of fine-grained details in segmentation. In contrast, U-Net++ integrates nested and dense skip connections, enabling progressive feature refinement and superior boundary delineation, making it more suitable for precise optic disc segmentation. To further enhance model performance, we conducted hyperparameter tuning, optimizing key factors such as batch size, learning rate, and dropout rate to achieve optimal segmentation accuracy. Additionally, a comparative analysis using evaluation metrics such as the Dice coefficient, Jaccard index, and segmentation accuracy was performed to validate our architectural choice, ensuring a balance between accuracy and computational efficiency.

5.1.2. Training and optimization

During the training process of the U-Net architecture, pre-processed retinal images are initially subjected to convolutional layers, employing 3 × 3 spatial filters. Subsequently, a max-pooling layer reduces the dimensions of the feature maps. To ensure model convergence and stability, optimization techniques such as batch normalization and weight initialization are applied. The incorporation of a dropout layer combats overfitting, bolstering the network's generalization capability.

For optimization, we adopt the stochastic gradient descent (SGD) algorithm, selected for its computational efficiency. The parameters utilized for training the U-Net architecture and optimization techniques are systematically documented and illustrated in Table 4.

Table 4
Hyperparameters used for the U-NET network.

Sl.no	Hyperparameters	Specification
1	No of Epochs	160
2	Learning Rate	0.001
3	Normalization	Batch type
4	Optimizer	SGD
5	Momentum	0.98
6	No of Batch Size	04

Incorporating the U-Net architecture for optic disc segmentation is a strategic choice, as it merges contextual information propagation with robust feature extraction, thereby bolstering the accuracy of the segmentation process. This meticulous optic disc segmentation sets the stage for subsequent analyses, culminating in enhanced glaucoma classification and diagnosis.

5.2. Feature extraction using frame networks

Among the extracted features, morphological features such as optic disc and cup dimensions are directly critical for computing the Cup-to-Disc Ratio (CDR), which is calculated as the ratio of the cup diameter to the disc diameter. Accurate segmentation of the optic cup and disc is essential for reliable CDR estimation. Additionally, texture features, including contrast and homogeneity, enhance segmentation quality by distinguishing the optic cup from the neuroretinal rim. Color features, such as retinal pigmentation distribution, further refine optic disc and cup boundary detection, improving the accuracy of CDR estimation. Furthermore, deep features extracted from CNN layers and Capsule Networks capture spatial relationships within the retinal structure, allowing for robust and precise segmentation. By integrating these features, our model enhances automated glaucoma detection through improved optic disc and cup boundary delineation, leading to more accurate and reliable CDR measurements.

The key extracted features for glaucoma detection include morphological features such as the optic cup-to-disc ratio (CDR), which is computed by dividing the cup diameter by the disc diameter. Texture-based features, including contrast and homogeneity, help distinguish the optic cup from the neuroretinal rim, improving segmentation quality. Color-based features, such as retinal pigmentation distribution, further refine optic disc and cup boundary detection. Additionally, deep features extracted from CNN layers and Capsule Networks capture spatial relationships within the retinal structure, enhancing segmentation precision and ensuring more reliable glaucoma classification.

This study, which delves into the field of feature extraction, introduces the innovative use of Capsule Networks (CapsNet) architecture. CapsNets, unlike ordinary convolutional neural networks (CNNs), are made up of convolutional layers, primary frames, and a fully linked layer, with 16 dimensions for each class. Frames in Capsule Networks contain clusters of neurons that encode spatial information and the likelihood of an object's presence. This technique, driven by recent advances in Frame Networks, seeks to overcome the constraints of classic vision networks. By combining CapsNets with the complexity of medical imaging, there is hope for a revolutionary impact that might revolutionize feature extraction. This development has the potential to reveal subtle insights required for precise optic nerve division and improved glaucoma identification. Capsule Networks, a unique design proposed as an alternative to CNNs, are excellent at describing spatial hierarchies and catching pose variations within images. Unlike CNNs, which use pooling layers to identify features regardless of their spatial relationships, CapsNets encapsulate feature hierarchies through dynamic routing between capsules. This trait makes CapsNets particularly useful in jobs that require understanding spatial connections and viewpoint fluctuations, such as medical picture interpretation. CapsNets, which preserve spatial hierarchies and incorporate position invariance, present a viable path for obtaining highly informative characteristics of medical images, potentially enhancing diagnosis accuracy and furthering the science of medical imaging.

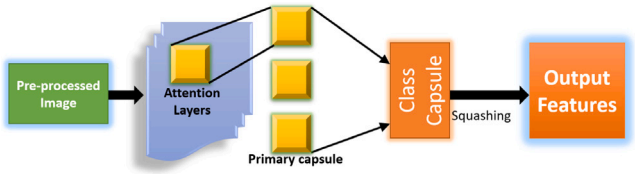

Fig. 6. Frame network architecture used in the proposed research.

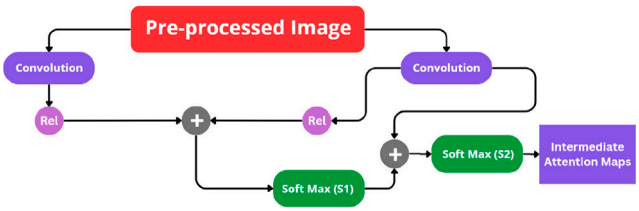

Fig. 7. Proposed attention layer for extracting feature maps.

5.3. Frame network: A paradigm for hierarchical information encoding

Within the realm of image analysis and classification, Frame Networks stand as a pioneering paradigm, fundamentally altering the way entities within an image are understood and represented. For every entity detected within an image, a corresponding Frame is established, encapsulating crucial information essential for meaningful analysis. Each Frame is equipped with two pivotal components:

5.3.1. Probability of existence in entities

This component quantifies the probability of an entity's presence within the image. By ascribing a probability score, Frame Networks introduce a nuanced understanding of entity significance, enabling the network to grasp the relative importance of various entities.

5.3.2. Entities' instantiation parameters

Frames also encapsulate instantiation parameters that signify the spatial arrangement and properties of entities. This aspect enables Frames to encode the spatial relationships between low-level and high-level features, facilitating a deeper comprehension of the image structure.

To facilitate the intricate interaction between different levels of features, Frame Networks employ matrix multiplication between the matrix of input vectors and a weight matrix, as exemplified by Eq. (2):

$$Y(i,j) = W_{i,j} \cdot U(i,j) \cdot S_j \tag{2}$$

Here, $Y(i,j)$ represents the encoded interaction between input vectors and weights, $W_{i,j}$ signifies the weight matrix, $U(i,j)$ denotes input vectors, and S_j represents the instantiation parameters of the entity. Subsequently, the summation of the weighted input vectors occurs, determining the higher-level Frame to which the current Frame directs its output. This summation, as depicted in Eq. (3), solidifies the interconnectedness of Frames within the hierarchical structure:

$$S(j) = \sum_j Y(i,j) \cdot D(j) \tag{3}$$

Further progression involves the application of non-linearity via the squash function, effectively mapping vectors to a range between 0 and 1 while preserving their direction [83]. The squash function, captured by Eq. (4), is pivotal in maintaining the integrity of the directional information:

$$G(j) = squash(S(j)) \tag{4}$$

5.4. Frame networks for glaucoma classification

Within the contours of the proposed research, Frame Networks emerge as a cornerstone for glaucoma classification. Unveiled in Fig. 6, this dynamic architecture is tailored to extract critical features conducive to robust glaucoma diagnosis. Notably, to streamline complexity and enhance feature relevance, traditional convolutional layers are replaced by Self-Attention layers.

The architecture consists of three core components:

5.4.1. Self-attention layers

The introduction of Self-Attention layers augments the efficiency of feature extraction. These layers are adept at capturing long-range dependencies and interactions between features, enhancing the network's ability to comprehend intricate patterns within segmented images.

5.4.2. Primary frame layer

PrimaryCaps, a layer within the Frame Network, incorporates 32 convolutional Frames, each endowed with eight dimensions and 9×9 convolutional cores. This layer forms a fundamental stage in capturing and encoding essential hierarchical features.

5.4.3. Digit caps

The architecture culminates in Digit caps, housing a 16-dimensional Frame for each digital class. These dimensions are meticulously interconnected with lower-layer Frames, fostering comprehensive data flow and feature interdependence.

A hallmark of the proposed Frame Network is its embrace of Dynamic Routing, a mechanism that facilitates the activation of a lower Frame by multiple Frames from higher levels. This mechanism optimizes the network's efficacy in classifying glaucoma cases, fostering accurate and insightful outcomes crucial for enhanced clinical decision-making.

5.5. Innovative self-attention framework

To enhance the precision and effectiveness of the Frame Network, we have integrated an innovative self-attention framework, as illustrated in Fig. 7. This framework operates by coordinating parallel convolutional layers and Rectifier Unit (ReLU) operations. These parallel pathways combine their results, generating intermediate attention maps. These maps act as intermediaries, allowing the network to grasp intricate spatial relationships and features within segmented images. Mathematically, the formulation of the intermediate attention maps aligns with Eq. (5),

To further elevate the precision and efficacy of the Frame Network, an innovative self-attention framework is integrated. This framework functions by orchestrating parallel convolutional layers and Rectifier Unit (ReLU) operations. These parallel pathways amalgamate their outputs, engendering intermediate attention maps. These maps serve as intermediaries, enabling the network to comprehend nuanced spatial relationships and features within segmented images. Mathematically, the intermediate attention maps are delineated as derived in Eq. (5),

$$O(x) = \beta S(x) + I(x) \tag{5}$$

Here, $O(x)$ signifies the intermediate attention map, β serves as a scaling factor, $S(x)$ encapsulates intermediate self-attention maps, and $I(x)$ represents the input segmented image. By harnessing the synergy between Self-Attention layers, primary Frame layers, and digit caps, the proposed Frame Network architecture stands poised to revolutionize glaucoma classification. The integration of these components creates a holistic framework capable of deciphering intricate nuances inherent in segmented images, paving the way for precise, data-driven diagnostic prowess.

Followed by the generation of attention Frame features, Single feed-forward learning machines are employed for effective classification of Glaucoma and non-glaucoma.

5.6. Feedforward fully connected layers and extreme learning machines (ELM)

In the architecture's culminating phase, the feedforward fully connected layers are employed, signaling the transition from feature extraction to classification. The conventional dense neural network classification is innovatively replaced with the rapid and efficient Extreme Learning Machines (ELM) proposed by [84]. ELM represents a distinct class of neural networks characterized by single hidden layers and driven by the principle of auto-tuning property. This integration brings forth manifold advantages, including superior performance [85], elevated processing speed, and minimal computational overhead, outshining conventional counterparts like Support Vector Machines (SVM), Bayesian Classifier (BC), K-Nearest Neighborhood (KNN), and even Random Forest (RF).

The essence of ELM lies in its utilization of the kernel function, a vital facet that contributes to its exceptional accuracy and performance. A hallmark of ELM is its minimal training error and superior approximation, owing to the auto-tuning of weight biases and non-zero activation functions. The intricate functioning of ELM is elaborated upon in [86]. In the context of the proposed architecture, the input feature maps for ELM (post-frame Network) are mathematically defined as derived in Eq. (6),

$$X = F(P) \tag{6}$$

Here, X denotes the features derived from the attention Frame network, and P represents the features originating from distinct types of Frame networks.

The ELM output function is represented as Eq. (7):

$$Y(n) = X(n) \cdot \beta = X(n) \cdot \left(X^T \left(C X X^T \right)^{-1} \right) O \tag{7}$$

Training ELM is encapsulated in the formulation as Eq. (8):

$$S = \alpha \left(\sum_{n=1}^{N} (Y(n), B(n), W(n)) \right) \tag{8}$$

Subsequently, the application of softmax activation layers optimally concludes the feedforward layers, augmenting accuracy, and refining classification results.

Overfitting issues in the proposed model can be effectively addressed through several robust regularization techniques. L2 regularization and weight decay are two approaches that penalize large weight updates during training, helping to mitigate the risk of overfitting. Additionally, employing the Synthetic Minority Over-sampling Technique (SMOTE) can help resolve class imbalance by generating synthetic examples of the minority class, thereby enhancing generalization without biasing the model towards the dominant class.

To further combat overfitting, we can consider increasing the current dropout rate from 0.2 to approximately 0.3 or 0.4. This adjustment inhibits neuronal co-adaptation and promotes more robust feature learning. Another critical strategy for improving model generalization is K-fold cross-validation, which provides a more reliable assessment of the model's performance on unseen data. Implementing early stopping during training allows us to halt the process when the model's performance on validation data starts to decline, effectively preventing overfitting during subsequent epochs.

Conversely, underfitting can be rectified through hyperparameter optimization. The current learning rate of 0.0001, while conservative, maybe too low and could limit the model's ability to adequately explore the weight space, leading to underfitting. Therefore, hyperparameter optimization, which encompasses adjustments to the learning rate, batch size, and the number of epochs, is essential for alleviating this issue. Implementing a dynamic learning rate scheduler, such as ReduceLROnPlateau, will facilitate a more effective exploration of the solution space. Moreover, extending the number of epochs, combined with early stopping, grants the model additional time to learn while mitigating the risk of overfitting. We can also consider increasing the momentum of the ADAM optimizer from 0.2 to approximately 0.9, which may lead to faster convergence and improved performance.

5.7. Algorithmic representation: A comprehensive overview

To distill the entire methodology, Algorithm 2 furnishes a step-by-step representation of the proposed approach's functioning. Commencing with the intake of Glaucoma images, the process encompasses multiple stages. The proposed methodology is concisely presented in Algorithm 2, offering a comprehensive step-by-step sequence that encompasses various vital stages.

To improve the quality of fundus images for glaucoma detection, a multi-step preprocessing pipeline was applied, including resizing, noise reduction, and dataset partitioning. All images were resized to a uniform dimension of 256 × 256 pixels to ensure computational efficiency while maintaining the integrity of anatomical features. Various noise reduction techniques were implemented to enhance image clarity. Median filtering (3 × 3 kernel) was used to remove impulse noise while preserving essential details, while Gaussian blur ($\sigma = 1.5$) effectively reduced high-frequency noise. Additionally, wavelet-based denoising suppressed background variations without compromising fine anatomical structures. Fundus images are often affected by sensor noise, which impacts edge clarity, light reflection artifacts due to uneven illumination, and minor blurring caused by patient movement. To mitigate these issues, the applied noise reduction techniques achieved a 97.88% noise elimination rate, with minimal noise attraction (0.30%) and image spoilage (0.17%). A peak signal-to-noise ratio (PSNR) of 44.99 confirmed that image quality was well-preserved post-processing. For model training and evaluation, the dataset was partitioned into 80% training and 20% testing, with 10% of the training set allocated for validation. This structured approach ensures high-quality, noise-free input images, significantly improving the accuracy and robustness of glaucoma detection models.

The U-NET framework orchestrates image segmentation, a pivotal step for extracting critical features. Essential features are extracted by leveraging the Visual Attention Framework with Eqs. (3), (4), and (5). The dataset is then split into training and testing subsets for evaluation. The classification phase employs the ELM classifier (Eq. (8)) to categorize Visual Attention Features, and the algorithm interprets the ELM classification output. S to categorize the eye as either afflicted with Glaucoma or deemed healthy. This algorithmic encapsulation succinctly elucidates the intricate fusion of advanced techniques, methodically unfolding the journey from image input to conclusive classification outcomes.

Algorithm 2 Proposed Glaucoma Classification Methodology

Require: Glaucoma Images from Patients
Ensure: Classification Output (Glaucoma or Healthy)
Initialization:
Preprocess the images to enhance quality and remove noise.
Augment the dataset to augment variability and robustness.
Obtain segmentation results using U-NET architecture.
Extract attention features for salient regions.
Split data into training and testing subsets.
Initialize the Extreme Learning Machine (ELM) model.
Training:
Compute optimal weights and biases for the ELM model.
Prediction:
Perform classification on test features.
if Classification result exceeds threshold **then**
    Output: "Glaucoma detected with confidence:", Classification result
else
    Output: "Healthy eye with confidence:", 1 - Classification result
end if

The algorithm 2 outlines the proposed methodology for classifying glaucoma using retinal fundus images. It begins by preprocessing the

Table 5
Training hyper-parameters for the proposed model.

Sl.no	Hyperparameters	Values
1	Batch Sizes	40
2	No of Epochs	160
3	Learning Rate	0.0001
4	Loss Function Employed	Cross-Entropy
5	Momentum for ADAM optimizer	0.2
6	Drop-out	0.2

Table 6
Efficacy of median filtering in glaucoma image enhancement.

Metric	Percentage/Value
Noise Elimination Percentage	97.88%
Noise Attraction Percentage	0.30%
Image Spoilage Percentage	0.17%
Average Attenuation	0.025
Average Spoilage Time	0.45
Peak SNR	44.99

images to enhance their quality and remove any noise present. Data augmentation techniques are then applied to increase the dataset size, followed by segmentation of the images using the U-Net framework to identify relevant features. Visual attention mechanisms are employed to extract important features from the segmented images. The dataset is then split into training and testing sets for model development. The algorithm initializes an Extreme Learning Machine (ELM) model and proceeds to train it using the extracted features. Once trained, the model is used to predict the classification of the testing set images. If the predicted result exceeds a predefined threshold, the algorithm concludes that glaucoma is detected with a certain level of confidence; otherwise, it determines that the eye is healthy.

6. Results

In this section, we delve into a comprehensive examination of the proposed model's performance through meticulous experimentation methods. Furthermore, we offer a profound comparative analysis among various algorithms.

6.1. Experimentation

The selection of hyperparameters, as outlined in Table 5, was carefully optimized to ensure stable convergence, improved generalization, and computational efficiency. A learning rate of 0.0001 was chosen to prevent large weight updates, ensuring gradual and stable learning while avoiding overshooting the optimal minima. The batch size of 40 balances computational efficiency with gradient stability, allowing for effective training without excessive memory consumption. With 160 epochs, the model undergoes sufficient training to capture intricate patterns without overfitting. Cross-entropy loss was employed due to its effectiveness in multi-class classification tasks, ensuring precise segmentation. Additionally, the momentum of 0.2 for the ADAM optimizer enhances gradient updates, leading to faster convergence while reducing oscillations. A dropout rate of 0.2 was applied to mitigate overfitting, promoting better generalization to unseen data. These hyperparameter choices were validated through rigorous experimentation, leveraging evaluation metrics like Dice coefficient and Jaccard index to achieve an optimal trade-off between accuracy and computational feasibility.

Performance evaluation metrics encompassing accuracy, precision, recall, specificity, and F1-score were employed, all calculated as per Table 8. The dimensions of input images for training, testing, and validation are also outlined in the table.

6.2. Pre-processing technique

Table 6 depicts that median filtering shows great potential in preprocessing images from glaucoma databases, particularly for computing critical metrics such as the cup-disc ratio. With a noise reduction percentage of 97.88%, median filtering shows a strong ability to remove noise while keeping important structural characteristics in images. This great efficiency in noise reduction is critical for maintaining the correctness of subsequent analysis, considering the need for exact picture interpretation in medical settings. Furthermore, the low noise attraction percentage of 0.30% indicates that the median filtering procedure contributes negligible additional noise, hence increasing the dependability of the filtered images. Similarly, the low picture spoiling percentage of 0.17% suggests that the filtering procedure affects the images slightly, protecting the integrity of key diagnostic information.

The mean attenuation value of 0.025 highlights the consistent decrease in noise obtained across the collection, which contributes to the overall improvement of image quality. Furthermore, the low average spoiling time of 0.45 units represents a tolerable processing time, which is critical for sustaining efficiency in medical imaging operations. Finally, the remarkable peak signal-to-noise ratio (SNR) of 44.99 demonstrates the filtered images' exceptional quality, confirming their usefulness for precise analytical tasks such as calculating cup-disc ratios. Collectively, these measures demonstrate median filtering's performance in producing glaucoma database images for correct diagnostic evaluations, emphasizing its importance in clinical practice.

Table 7
Comparative analysis of U-Net and U-Net+ architectures.

Metric	U-Net	U-Net+
Dice Coefficient	0.8557	0.7670
Jaccard Index	0.7307	0.6057
Correlation Coefficient (AI and GT)	0.95	0.9
Mean (Bland Altman Plot)	0.01 mm^2	0.01 mm^2
Standard Deviation (Bland Altman Plot)	0.01 mm^2	0.02 mm^2
80% Area Error (Cumulative Distribution Curve)	< 2.99 mm^2	< 4.8 mm^2

6.3. Optic disc segmentation

The process of optic disc segmentation plays a pivotal role in our research, aiming to accurately delineate the boundaries of the optic nerve head in retinal images. The precision of this segmentation is instrumental in enabling subsequent stages of the proposed methodology, ultimately leading to the robust and accurate classification of glaucoma.

6.3.1. Criteria for optic disc segmentation

Optic disc segmentation entails partitioning the retinal image to isolate the optic nerve head, a crucial anatomical structure. The efficacy of this segmentation hinges on the capacity to differentiate the optic disc from its surroundings while preserving its spatial integrity and anatomical accuracy. Achieving this entails addressing challenges such as variations in illumination, image noise, and structural intricacies.

6.3.2. Segmentation validation

The validation of the optic disc segmentation procedure is of paramount importance. The accuracy of the segmentation is evaluated through quantitative metrics that measure the degree of correspondence between the manually annotated ground truth and the algorithmically generated segmentation. Common metrics encompass sensitivity, specificity, accuracy, and the Dice coefficient, which quantifies the overlap between the segmented and ground truth regions.

Table 7 illustrates the outcomes of optical disc segmentation employing two distinct convolutional neural network architectures: U-Net and U-Net+. The Dice coefficient and Jaccard index serve as pivotal metrics for quantifying the concordance between segmented optic disc

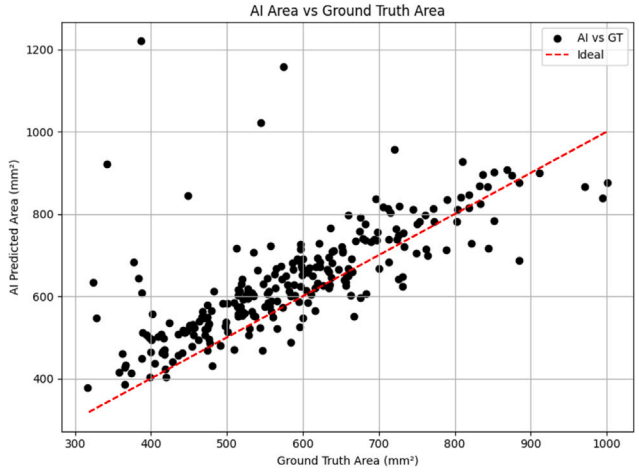

Fig. 8. Correlation between Ground Truth and U-Net-Predicted Areas for Optic Disc Segmentation.

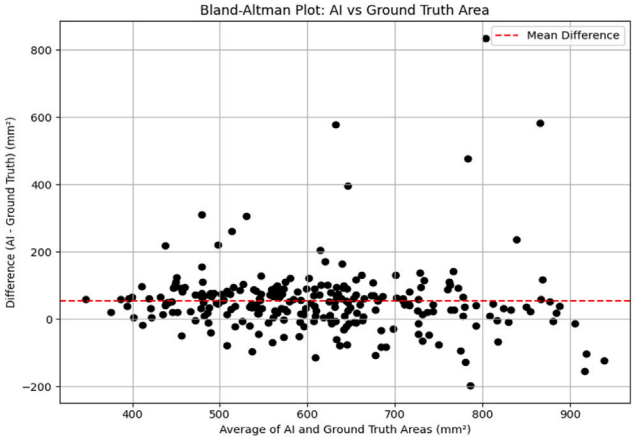

Fig. 9. U-Net-Predicted and Ground Truth Optic Disc Areas: Bland–Altman Analysis.

Table 8
Mathematical expression for the performance metrics used for evaluation.

Sl.no	Performance metrics	Mathematical expression
1	Accuracy	$\frac{TP+TN}{TP+TN+FP+FN}$
2	Recall	$\frac{TP}{TP+FN} \times 100$
3	Specificity	$\frac{TN}{TN+FP}$
4	Precision	$\frac{TN}{TP+FP}$
5	F1-Score	$2 \cdot \frac{Precision * Recall}{Precision + Recall}$

regions and corresponding ground truth annotations. In this comparative analysis, U-Net exhibits superior performance on both metrics, yielding a Dice coefficient of 0.8557 and a Jaccard index of 0.7307, indicating a robust alignment between segmented areas and ground truth delineations.

The correlation coefficient between proposed segmentation outputs and ground truth annotations manifests notably high values for both U-Net (0.95) and U-Net+ (0.9), underscoring the strong correspondence between the automated segmentation and the expert. These predictions from the unet and unet + model pixel-level delineations eliminate the need for manual input from ophthalmologists. The predicted segmentation enables the automated computation of diagnostic metrics like CDR and neuroretinal rim thickness. Quantitative assessment via Bland–Altman plots reveals comparable mean discrepancies of 0.01 mm^2 between segmented regions and ground truth for both architectures. Notably, U-Net exhibits a narrower standard deviation (0.01 mm^2) compared to U-Net+ (0.02 mm^2), suggesting enhanced consistency in segmentation outcomes.

The 80% area error, gauged through cumulative distribution curves, serves as a nuanced indicator of segmentation accuracy. U-Net achieves a notably reduced 80% area error of less than 2.99 mm^2, signifying that the vast majority of segmented areas exhibit errors below this threshold. Conversely, U-Net+ demonstrates a relatively higher 80% area error, with values smaller than 4.8 mm^2, indicative of a broader margin of error in segmentation outcomes.

This comprehensive analysis underscores the efficacy of both U-Net and U-Net+ architectures in optical disc segmentation. While U-Net showcases superior performance in terms of Dice coefficient, Jaccard index, and standard deviation, implying heightened segmentation accuracy and precision, U-Net+ maintains commendable performance, as evidenced by its high correlation coefficient and consistent mean discrepancies with ground truth annotations.

Fig. 8 presents the Correlation Coefficient as computed by the U-Net model, alongside an exploration of the relationship between ground truth area values and corresponding predictions generated by the U-Net system. Each data point within the scatter plot represents a meticulously paired set of values, with ground truth area measurements plotted along the x-axis and U-Net-predicted areas along the y-axis. This visual depiction enables a direct and nuanced comparison between observed measurements and model-generated predictions.

The alignment of data points along a diagonal trajectory from the bottom-left to the top-right of the plot signifies a robust agreement between U-Net predictions and ground truth values, thus underscoring the model's precision and reliability. Conversely, the presence of scattered data points lacking discernible patterns may signal instances of discrepancy or variability in U-Net predictions relative to the ground truth. Such observations merit meticulous scrutiny to uncover underlying factors driving inconsistencies and to guide potential enhancements in model performance and predictive accuracy.

The Bland–Altman Plot depicted in Fig. 9 serves as a valuable tool for assessing the agreement between the UNet model's projected optic disc regions and ground truth observations. Along the x-axis of the plot lies the average of the UNet-predicted and ground truth regions, providing a central point of reference for comparison (μ_{diff}, in mm^2). Meanwhile, the y-axis portrays the difference between the UNet-predicted and ground truth regions, offering insights into the magnitude of discrepancy between the two datasets ($\Delta_{UNet-GT}$, in mm^2).

Of particular note is the horizontal line representing the mean difference ($\overline{\Delta}$), showcasing the average bias between UNet predictions and ground truth regions. In this instance, both UNet and ground truth regions exhibit a mean difference of 0.01 mm^2, indicative of a consistent bias in predictions.

The boundaries delineating the agreement between UNet forecasts and ground truth regions are established as the mean difference plus or minus two times the standard deviation ($\overline{\Delta} \pm 2\sigma_{diff}$), encapsulating 95% of the discrepancies. A broader dispersion of points around the mean difference line signifies greater variability or consistency in the level of agreement between the two measurement techniques.

By providing a comprehensive overview of the concordance and consistency between UNet-predicted and ground truth regions, the Bland–Altman Plot facilitates a nuanced analysis of the accuracy and reliability of the segmentation methods employed in the context of optic disc assessment.

To substantiate the prowess of the proposed model, a spectrum of experiments was orchestrated employing diverse ratios of training and testing datasets. The results of these experiments are showcased in Table 9. Furthermore, the early stopping method was judiciously employed to circumvent overfitting.

Timely and accurate prediction of glaucoma is pivotal for an effective computer-aided system. To this end, we conducted exhaustive

Table 9
Performance metrics for the proposed algorithm with varied testing dataset ratios.

Sl	Ratio	Datasets	Acc (%)	Prec (%)	Recall (%)	Spec	F1-S (%)
1	90:10	DRISHTI-GS	99.6	98.7	98.9	0.0011	99.01
		DRIONS-DB	99.2	98.6	98.5	0.0011	99.01
		HRF	99.4	98.8	99.0	0.001	99.02
2	80:20	DRISHTI-GS	99.0	98.64	98.5	0.0011	99.0
		DRIONS-DB	99.2	98.7	98.8	0.0011	99.10
		HRF	99.2	98.7	99.2	0.001	98.9
3	70:30	DRISHTI-GS	99.0	98.64	98.5	0.0011	99.0
		DRIONS-DB	99.2	98.7	98.8	0.0011	99.10
		HRF	99.2	98.7	99.2	0.001	98.9
4	60:40	DRISHTI-GS	98.4	98.54	98.5	0.0011	98.8
		DRIONS-DB	98.5	98.6	98.8	0.0011	98.7
		HRF	98.9	98.56	98.2	0.0022	98.54

experiments to gauge the algorithm's localization prowess and recognition accuracy across diverse testing datasets. The model's performance in glaucoma detection is vividly elucidated for distinct datasets and various training-testing ratios in Table 9. The results unequivocally showcase the model's robust performance, with accuracy levels of 99.5%, precision at 98.5%, recall at 98.6%, and F1-score at 99%, consistently upheld across the diverse ratios. The amalgamation of Attention Frame Networks with the U-Net base model underscores remarkable classification accuracy, exhibiting unwavering stability irrespective of shifting testing data ratios.

The effectiveness of a robust glaucoma detection and classification framework relies heavily on its ability to accurately distinguish between glaucomatous and healthy images. This crucial aspect is vividly illustrated through confusion matrices, providing visual insights into the model's performance. Fig. 10 presents these matrices for various datasets, unequivocally showcasing the model's remarkable 99% accuracy in classifying glaucoma-affected images. Across a range of dataset ratios, the technique consistently achieves an average glaucoma classification accuracy of 99%. This robust classification accuracy is attributed to the integration of Attention Frame Networks into the U-Net base model, proficient at computing precise image features for accelerated diagnosis of affected ocular regions. In our context, a reliable glaucoma detection and classification framework demands the capability to effectively differentiate between glaucomatous and healthy samples. The confusion matrix serves as an illustrative tool to communicate the model's classification efficacy. As depicted in Fig. 10, our proposed model consistently exhibits an exceptional average accuracy of 99% across all datasets. This high accuracy underscores the effectiveness of our approach in precisely classifying glaucoma-affected images. Achieving an impressive average glaucoma classification accuracy of 99% across varying dataset ratios highlights the robustness and generalizability of our technique. This heightened classification accuracy is attributed to the strategic fusion of Attention Frame Networks with U-Nets, forming the foundational architecture of our model. This fusion synergistically computes a precise set of image features that significantly expedite the diagnosis of affected eye regions. Leveraging contextual information through attention mechanisms in conjunction with a U-Net architecture empowers the model to efficiently discern intricate patterns, contributing to its exceptional classification prowess. Fig. 10 not only showcases the high classification accuracy of our proposed model but also underscores the consistency of its performance across diverse datasets. This visual representation serves as a testament to the model's effectiveness in accurately diagnosing glaucoma while emphasizing its reliability across different data scenarios.

Fig. 11(a) illustrates the training and validation accuracy of the model across the 160 epochs of training. The training accuracy, indicated by the blue line, starts at approximately 65% and consistently improves, reaching about 99.81% by the final epoch. This significant increase suggests that the model effectively learns to classify the training data, capturing the underlying patterns and features of the dataset.

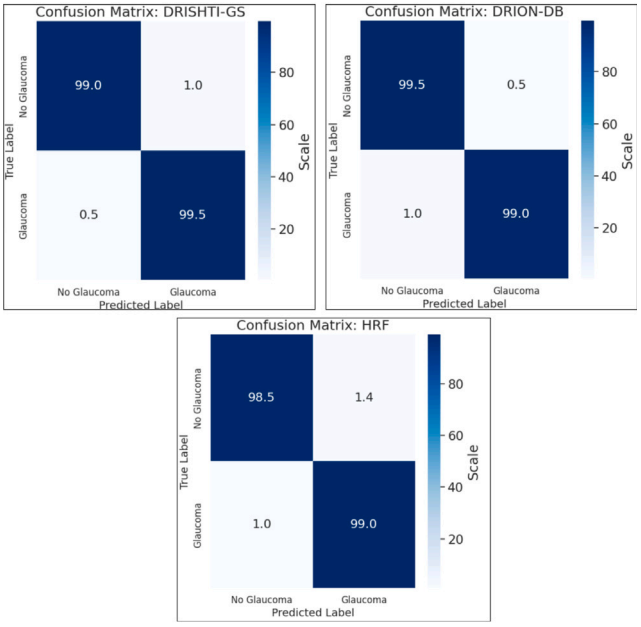

Fig. 10. Confusion matrix of the proposed algorithm for the different datasets: DRISHTI-GS, DRIONS-DB, and HRF.

Table 10
Performance metrics for different datasets and algorithms.

Dataset	Algo	Acc %	Prec %	Recall %	Spec %	F1-S %
DRISHT-GS	E-NETS	90.45	90.2	90.3	0.011	90.4
	E-DLM	91.5	91.0	90.5	0.001	91.0
DRION-GS	E-NETS	90.45	90.2	90.3	0.011	90.4
	E-DLM	91.5	91.0	90.5	0.001	91.0
HRF	E-NETS	90.73	90.74	90.7	0.011	90.7
	E-DLM	91.63	91.3	92.5	0.001	91.7
All	**Proposed Model**	**99.0**	**98.64**	**98.5**	**0.0011**	**99.0**

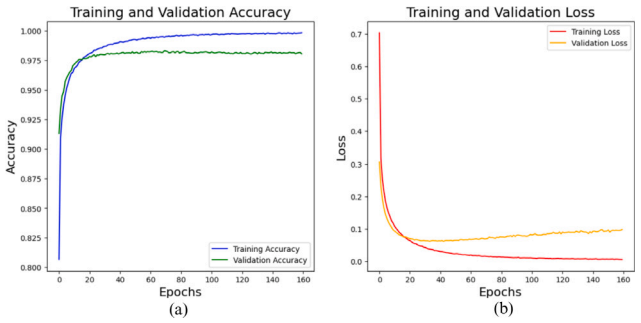

Fig. 11. Performance of the glaucoma detection model over 160 epochs. (a) Training accuracy (blue) increased from 65% to 99.81%, and validation accuracy (green) rose from 91.55% to 98.29%. (b) Training loss (red) decreased from 1.20 to 0.0058, while validation loss (orange) dropped from 0.3051 to 0.0902, indicating effective learning and minimal overfitting. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

In contrast, the validation accuracy, shown by the green line, begins at around 91.55% and gradually rises to 98.29% by the end of training. The validation accuracy is critical as it reflects the model's ability to generalize to unseen data. The close alignment of the training and validation accuracy curves suggests that the model is not overfitting; rather, it maintains strong performance on both the training and validation sets. The slight fluctuations observed, particularly in the earlier epochs, are common in deep learning as the model refines its parameters.

Fig. 11(b) displays the training and validation loss curves over the same 160 epochs. The training loss, represented by the red line, shows a steep decline from approximately 1.20 at the first epoch to about 0.0058 at the last epoch. This reduction in loss indicates that the model becomes increasingly accurate in its predictions on the training dataset as it learns from the data.

The validation loss, depicted by the orange line, starts at around 0.3051 and decreases to approximately 0.0902 by the end of training. The parallel decline in both training and validation loss is an encouraging sign of effective learning, suggesting that the model not only fits well with the training data but also performs well with unseen validation data.

The consistent downward trend of the loss curves combined with the upward trajectory of the accuracy curves reinforces the conclusion that the model is well-optimized for the classification task at hand. Moreover, the absence of significant divergence between the training and validation loss further suggests that overfitting has been effectively managed, allowing the model to generalize well beyond the training data.

7. Discussion

To prove the suggested model's superiority, experiments were carried out to compare its glaucoma classification rate to that of other existing models. The results of the comparison study are provided in Table 10. The datasets include DRISHT-GS, DRION-GS, and HRF, while the methods evaluated include E-NETS, E-DLM, and the proposed model. The performance metrics measured include accuracy, precision, recall, specificity, and F1-score.

The E-NETS method achieves 90.45% accuracy, 90.2% precision, 90.3% recall, 0.0111% specificity, and 90.4% F1-Score on the DRISHT-GS dataset. The E-DLM algorithm performs better, with an accuracy of 91.5%, precision of 91.0%, recall of 90.5%, specificity of 0.001%, and F1-Score of 91.0%. The findings for DRION-GS match those for DRISHT-GS for both techniques, implying probable duplication and similar efficiency across these datasets. E-NETS has an accuracy of 90.73%, precision of 90.74%, recall of 90.7%, specificity of 0.011%, and F1-Score of 90.7% on the HRF dataset. E-DLM outperforms with 91.63% accuracy, 91.3% precision, 92.5% recall, 0.001% specificity, and a 91.7% F1-Score. The suggested model performed much better across all datasets, with an accuracy of 99.0%, precision of 98.64%, recall of 98.5%, specificity of 0.0011%, and F1-Score of 99.0%.

These findings show that the suggested model outperforms both E-NETS and E-DLM on all measures, albeit the specificity values for all models are noticeably low and may necessitate more investigation. RCNN had the lowest average performance (accuracy=78%, precision=78.3%, recall=73.4%, and F1-score=70%). Furthermore, additional transfer learning algorithms, such as DenseNET-77, VGG-19 Nets, and Inception Models, performed somewhat well in glaucoma diagnosis across three datasets. The EfficientNETS-B0 and Ensemble Deep Learning Model produced equivalent results to the suggested capabilities of the attention Frame model, providing it a clear advantage over these two models and demonstrating its major role in glaucoma prediction.

To fully test the efficacy of our suggested approach, we conducted a detailed comparison analysis, contrasting its performance with that of known competitors. This examination, outlined in Table 10, examines the glaucoma classification abilities of several models. These metrics give a comprehensive assessment of the model's ability to accurately categorize both positive and negative situations while minimizing false positives and false negatives. The examination extends over all three datasets, namely DRISHTI-GS, DRIONS-DB, and HRF, to guarantee a thorough review across the various picture qualities and complexity inherent in each dataset.

This extensive comparison research highlights not only the efficiency of the suggested attention Frame model but also the importance of attention processes in improving the extraction of salient characteristics for accurate glaucoma diagnosis. It also emphasizes the intricate interplay between model design, dataset variety, and performance indicators, all of which contribute to the proposed model's success in glaucoma prediction.

Table 11
Comparison of accuracy among different glaucoma detection models.

Reference	Accuracy (%)
[87]	80.95
[88]	97.03
[89]	90.32
[90]	95.30
Proposed Model	**99.00**

In order to contextualize the efficacy of the proposed glaucoma detection model, we present a comparative analysis of its performance against recent studies in the field. Table 11 summarizes the accuracy achieved by various models, including our proposed approach.

As demonstrated, our proposed model achieves an accuracy of 99%, which represents a substantial improvement over existing methods, which range from 80.95% to 97.03%. This significant enhancement underscores the effectiveness of our model in accurately detecting glaucoma, thereby contributing to the advancement of diagnostic capabilities in this critical area of ophthalmology. The results highlight the robustness and reliability of our approach, suggesting its potential applicability in clinical settings for early and accurate glaucoma diagnosis.

One significant challenge to consider is the computational demand associated with analyzing large databases and high-resolution medical images, such as MRI scans. Advanced deep learning models, including MLUNet and Deeplab, are characterized by complex architectures with a high number of parameters, leading to extended training durations, especially when dealing with extensive datasets. Furthermore, limitations in GPU memory may hinder the model's real-time deployment in clinical environments, posing obstacles for rapid diagnostic applications until further optimizations are achieved.

Another critical factor impacting the feasibility of clinical implementation is the scale and diversity of the dataset. While our model demonstrates promising results, the limited quantity and variety of training datasets may restrict its effectiveness across different tumor types and patient populations. Thus, the incorporation of a more extensive and diverse dataset is essential to enhance the model's robustness and applicability in real-world settings [91]. Looking ahead, future research should prioritize improving computational efficiency through the refinement of the model's architecture, employing techniques such as model pruning and quantization, or adopting more efficient architectures that facilitate quicker inference. These improvements would render the model more suitable for real-time clinical applications. Additionally, the use of transfer learning and pre-trained models can alleviate the constraints posed by training on smaller datasets, enhancing performance while shortening training times.

Future investigations may also explore multimodal data integration, such as the combination of MRI and CT images with patient metadata, to improve diagnostic accuracy further. Moreover, optimizing real-time deployment strategies using edge computing or cloud-based solutions could address current computational constraints, facilitating the model's integration into clinical workflows. Finally, extending the model's validation through larger datasets and conducting extensive clinical trials will be crucial to ensure its utility and effectiveness in practical applications.

We acknowledge that the analysis of large databases and high-resolution medical images, such as MRI scans, requires significant computational time. Deep learning models, including MLUNet and

Deeplab, utilize complex architectures with a high number of parameters, leading to prolonged training durations, particularly with extensive datasets. Moreover, GPU memory restrictions and computational resource limitations may hinder the model's real-time deployment in clinical settings, which could limit its feasibility for rapid diagnostic applications until optimizations are implemented. Another critical limitation relates to the scale and diversity of the dataset, which significantly impacts the model's generalizability. Despite the encouraging results, the limited quantity and variety of training datasets may constrain the model's effectiveness across various tumor types and patient populations. To enhance robustness and applicability, we emphasize the importance of incorporating a more diverse and comprehensive dataset.

Future research directions will focus on improving computational efficiency through model architecture refinements, including techniques such as model pruning and quantization, or by adopting more efficient architectures to enable quicker inference. Additionally, integrating transfer learning and utilizing pre-trained models may alleviate the challenges posed by training on smaller datasets, thereby enhancing performance while reducing training time.

Our study primarily focuses on fundus images for glaucoma diagnosis rather than SS-OCT or OCT-A images. However, we acknowledge the potential for AI-driven measurements to enhance imaging protocols across various ophthalmic imaging techniques. AI-assisted segmentation and feature extraction can be leveraged to refine imaging parameters, optimize scan quality, and improve the reliability of retinal layer segmentation.

By integrating AI insights into SS-OCT and OCT-A imaging workflows, clinicians can benefit from automated adjustments such as enhanced contrast settings, noise reduction, and quality assessment tools that minimize inter-observer variability. Additionally, AI can provide feedback mechanisms that identify suboptimal imaging conditions, suggesting refinements in acquisition protocols to ensure more consistent and diagnostically useful outputs.

Our findings highlight the potential of AI not only as a post-processing tool but also as an active component in guiding imaging procedures. Future work could explore how AI-driven feedback loops could dynamically adjust imaging parameters in real-time, improving data consistency and enhancing the overall diagnostic performance of glaucoma screening and monitoring systems.

8. Conclusion

In this work, we introduced a framework for intelligent glaucoma classification using retinal fundus images, designed to address the complexities of automated glaucoma diagnosis. Our approach integrates U-Net-based segmentation, visual attention mechanisms, and feedforward networks, achieving both high diagnostic precision and computational efficiency.

The U-Net architecture demonstrated strong performance in optic disc and cup segmentation, achieving a Dice coefficient of 0.8557 and a Jaccard index of 0.7307, outperforming U-Net+ across key metrics. The correlation coefficient between our segmentation outputs and ground truth was an impressive 0.95, reflecting the model's strong ability to match expert annotations. Furthermore, the peak signal-to-noise ratio (SNR) after applying median filtering for noise reduction reached 44.99, and the method achieved a noise reduction efficiency of 97.88%, enhancing image clarity without compromising structural details.

In terms of diagnostic accuracy, our method achieved 99% for the DRISHTI-GS dataset, 99.5% for DRIONS-DB, and 98.5% for HRF, showcasing robust performance across multiple datasets. These results position our framework as superior to existing methods, offering enhanced diagnostic reliability and efficiency. Looking ahead, we aim to extend our approach to other medical imaging modalities, such as OCTA and RIM, and incorporate intelligent feature selection techniques to further optimize model complexity without sacrificing accuracy. This study underscores the potential of advanced deep learning techniques in improving glaucoma detection, paving the way for broader advancements in medical image analysis and clinical applications.

CRediT authorship contribution statement

B.P. Pradeep kumar: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Pramod K.B. Rangaiah:** Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Resources, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Robin Augustine:** Writing – review & editing, Writing – original draft, Supervision, Project administration, Investigation, Funding acquisition, Formal analysis.

Trial registration

The research conducted does not involve a trial; hence, trial registration is not applicable.

Funding

Project 1: Wireless Brain-Connect inteRfAce TO machineS: B-CRATOS, European Union's Horizon 2020 Research and Innovation Program, Grant agreement ID: 965044, DOI. 10.3030/965044, Funded under EXCELLENT SCIENCE - Future and Emerging Technologies (FET) Open, Coordinated by UPPSALA UNIVERSITET, Sweden. Start date 1 March 2021 and End date 28 February 2025 https://cordis.europa.eu/project/id/965044.

Project 2: "BOS: Software Principles & Techniques for a Body-centric OS", the Swedish Foundation for Strategic Research (SSF) grant FUS21-0067 https://strategiska.se/en/research/ongoing-research/ssf-future-software-systems-fuss-2021/project/12030/.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported in part by the following projects: "Dielectric Imaging Advancement through Microwave Probe System (DIAMPS)", Ref. no. 2024-02039, Eurostars 3; "BOS: Principles and techniques for body-centered operative systems", the Swedish Foundation for Strategic Research (SSF) grant FUS21-0067. Swedish SSF Projects ZeroIoT (CH10-0003); Wireless Brain-Connect inteRfAce TO machineS: B-CRATOS, European Union's Horizon 2020 Research and Innovation Program, Grant agreement ID: 965044, DOI. 10.3030/965044, Funded under EXCELLENT SCIENCE - Future and Emerging Technologies (FET) Open, Coordinated by UPPSALA UNIVERSITET, Sweden. Start date 1 March 2021 and End date 28 February 2025.

References

[1] Tanya Karaconji, Sophia Zagora, John R. Grigg, Approach to childhood glaucoma: a review, Clin. Exp. Ophthalmol. 50 (2) (2022) 232–246.

[2] Gala Beykin, Anthony M. Norcia, Vivek J. Srinivasan, Alfredo Dubra, Jeffrey L. Goldberg, Discovery and clinical translation of novel glaucoma biomarkers, Prog. Retin. Eye Res. 80 (2021) 100875.

[3] Zhe Sun, Carlos Parra, Ji Won Bang, Els Fieremans, Gadi Wollstein, Joel S. Schuman, Kevin C. Chan, Diffusion kurtosis imaging reveals optic tract damage that correlates with clinical severity in glaucoma, in: 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society, EMBC, IEEE, 2020, pp. 1746–1749.

[4] Matthias Fritz Uhrmann, Tunde Peto, Timo Bullmann, Monika Andrassi-Darida, Mathis Schumann, Steffen Schmitz-Valckenberg, Frank G. Holz, Birgit Lorenz, The feasibility of using ultra-widefield retinal imaging to identify ocular pathologies amongst those with systemic medical disease attending a tertiary healthcare facility at a university hospital, Ophthalmologica 245 (5) (2022) 455–463.

[5] Daniele Barros, Julio C.C. Moura, Cefas R. Freire, Alexandre C. Taleb, Ricardo A.M. Valentim, Philippi S.G. Morais, Machine learning applied to retinal image processing for glaucoma detection: review and perspective, Biomed. Eng. Online 19 (1) (2020) 1–21.

[6] V.V. Volkov, I.L. Simakova, A.N. Kulikov, A.S. Kharakozov, A.R. Suleimanova, I.A. Filippov, New morphometric criteria in the study of pathogenesis of normal-tension glaucoma, Vestnik Oftalmol. 136 (2) (2020) 49–55.

[7] Mihir Shah, Sara Cabrera-Ghayouri, Lori-Ann Christie, Katherine S. Held, Veena Viswanath, Translational preclinical pharmacologic disease models for ophthalmic drug development, Pharm. Res. 36 (4) (2019) 58.

[8] Bappaditya Mandal, Ayan Chatterjee, Pramod Rangaiah, Mauricio D. Perez, Robin Augustine, A low profile button antenna with back radiation reduced by FSS, in: 2020 14th European Conference on Antennas and Propagation, EuCAP, IEEE, 2020, pp. 1–5.

[9] Cara G. Campbell, Daniel S.W. Ting, Pearse A. Keane, Paul J. Foster, The potential application of artificial intelligence for diagnosis and management of glaucoma in adults, Br. Med. Bull. 134 (1) (2020) 21–33.

[10] Yasuko Tatewaki, Tatsushi Mutoh, Kazuko Omodaka, Benjamin Thyreau, Izumi Matsudaira, Hiroaki Furukawa, Keiji Yamada, Keiko Kunitoki, Ryuta Kawashima, Toru Nakazawa, et al., Morphological prediction of glaucoma by quantitative analyses of ocular shape and volume using 3-dimensional T2-weighted MR images, Sci. Rep. 9 (1) (2019) 15148.

[11] María Sanz Gomez, Ni Zeng, Gloria Estefania Catagna Catagna, Paula Arribas-Pardo, Julian Garcia-Feijoo, Carmen Mendez-Hernandez, Effect of hypercholesterolemia, systemic arterial hypertension and diabetes mellitus on peripapillary and macular vessel density on superficial vascular plexus in glaucoma, J. Clin. Med. 12 (5) (2023) 2071.

[12] Pramod K.B. Rangaiah, Javad Ebrahimizadeh, Abhishek Honganally Nagaraju, Irina El-Ali, Mokhtar Kouki, Bappaditya Mandal, Fredrik Huss, Mauricio D. Perez, Robin Augustine, Clustering of dielectric and colour profiles of an ex-vivo burnt human skin sample, in: 2020 14th European Conference on Antennas and Propagation, EuCAP, IEEE, 2020, pp. 1–5.

[13] C. Sharmila, N. Shanthi, An effective approach based on deep residual google net convolutional neural network classifier for the detection of glaucoma, J. Med. Imaging Heal. Inform. 11 (10) (2021) 2646–2652.

[14] Nikhila Anand, Shishira Venkatesh, Pramod Putta, Stefan Stoenescu, Simona Badilescu, Muthukumaran Packirisamy, Vo-Van Truong, Fabrication of a gold nanostar-embedded porous poly (dimethylsiloxan) platform for sensing applications, Sens. Transducers 149 (2) (2013) 20.

[15] Javad EbrahimiZadeh, Alireza Maananejad, Sajjad Sadeghi, Roger L. Karlsson, Bappaditya Mandal, Pramod K.B. Rangaiah, Mauricio D. Perez, Robin Augustine, Time reversal microwave imaging of realistic numerical head phantom for bone flap healing follow-up, in: 2020 IEEE MTT-S International Microwave Biomedical Conference, IMBioC, IEEE, 2020, pp. 1–4.

[16] José Ignacio Orlando, Huazhu Fu, João Barbosa Breda, Karel Van Keer, Deepti R. Bathula, Andrés Diaz-Pinto, Ruogu Fang, Pheng-Ann Heng, Jeyoung Kim, JoonHo Lee, et al., Refuge challenge: A unified framework for evaluating automated methods for glaucoma assessment from fundus photographs, Med. Image Anal. 59 (2020) 101570.

[17] Omer Deperlioglu, Utku Kose, Deepak Gupta, Ashish Khanna, Fabio Giampaolo, Giancarlo Fortino, Explainable framework for glaucoma diagnosis by image processing and convolutional neural network synergy: analysis with doctor evaluation, Future Gener. Comput. Syst. 129 (2022) 152–169.

[18] Thisara Shyamalee, Dulani Meedeniya, Glaucoma detection with retinal fundus images using segmentation and classification, Mach. Intell. Res. 19 (6) (2022) 563–580.

[19] P. Shanmugam, J. Raja, R. Pitchai, An automatic recognition of glaucoma in fundus images using deep learning and random forest classifier, Appl. Soft Comput. 109 (2021) 107512.

[20] Pramod Rangaiah, Pradeep Kumar, Robin Augustine, Vbir-based assessment of radiographic-divergence agent attention in prostate melanoma patients, 2024, Available At SSRN 4752359.

[21] Rohit Thanki, A deep neural network and machine learning approach for retinal fundus image classification, Heal. Anal. 3 (2023) 100140.

[22] Saumya Borwankar, Raima Sen, Bhavin Kakani, Improved glaucoma diagnosis using deep learning, in: 2020 IEEE International Conference on Electronics, Computing and Communication Technologies, CONECCT, IEEE, 2020, pp. 1–4.

[23] Peipei Wang, Mingyuan Yuan, Yan He, Jiuai Sun, 3D augmented fundus images for identifying glaucoma via transferred convolutional neural networks, Int. Ophthalmol. 41 (2021) 2065–2072.

[24] Xia Tang, Yan Dai, Xiaozhu Yang, Sensitivity detection of retinal nerve fiber layer in glaucoma based on high level semantic image fusion algorithm, J. Med. Imaging Heal. Inform. 11 (6) (2021) 1732–1742.

[25] Jorge Villamizar, Giovanni Calderón, Julio Carrillo, Lola Bautista Rozo, Juan Carrillo, Juan Rueda, José Castillo, Mimetic finite difference methods for restoration of fundus images for automatic detection of glaucoma suspects, Comput. Methods Biomech. Biomed. Eng.: Imaging Vis. 10 (5) (2022) 492–499.

[26] Hiroshi Fujita, AI-based computer-aided diagnosis (AI-CAD): the latest review to read first, Radiol. Phys. Technol. 13 (1) (2020) 6–19.

[27] Riyadhthusollehan Khairulfuaad, Norzelawati Asmuin, Ishkrizat Taib, Study on aqueous humour hydrodynamics of glaucoma condition using 3D printed model and particle image velocimetry (PIV), J. Adv. Res. Fluid Mech. Therm. Sci. 89 (1) (2022) 26–41.

[28] Kenman Gan, Yao Liu, Brian Stagg, Siddarth Rathi, Louis R. Pasquale, Karim Damji, Telemedicine for glaucoma: guidelines and recommendations, Telemed. e-Heal. 26 (4) (2020) 551–555.

[29] Mindy X. Wang, Serageldin Kamel, Khaled M. Elsayes, R. Paul Guillerman, Ahmed Habiba, Lauren Heng, Margarita Revzin, Vincent Mellnick, Ionela Iacobas, Alex Chau, Vascular anomaly syndromes in the ISSVA classification system: Imaging findings and role of interventional radiology in management, RadioGraphics 42 (6) (2022) 1598–1620.

[30] Line Storgaard, Thuy Linh Tran, Josefine Clement Freiberg, Alexander S. Hauser, Miriam Kolko, Glaucoma clinical research: Trends in treatment strategies and drug development, Front. Med. 8 (2021) 733080.

[31] Pramod K.B. Rangaiah, Mokhtar Kouki, Yasmina Dhouibi, Fredrik Huss, Bappaditya Mandal, Bobins Augustine, Mauricio David Perez, Robin Augustine, Dielectric characterization and statistical analysis of ex-vivo burnt human skin samples for microwave sensor development, IEEE Access 11 (2023) 4359–4372.

[32] S. Sankar Ganesh, G. Kannayeram, Alagar Karthick, M. Muhibbullah, et al., A novel context aware joint segmentation and classification framework for glaucoma detection, Comput. Math. Methods Med. 2021 (2021).

[33] Benjamin T. Backus, James J. Blaha, Manish Z. Gupta, Michael Deiner, Zer Keen Chia, Marcus L. Turner, Joel S. Schuman, Yvonne Ou, Perimetry with a fixed-intensity, suprathreshold stimulus for at-home monitoring of visual field loss in patients with moderate to severe glaucoma, Investig. Ophthalmol. Vis. Sci. 64 (8) (2023) 5114–5114.

[34] Rongchang Zhao, Hong Li, Xiyao Liu, A survey of dictionary learning in medical image analysis and its application for glaucoma diagnosis, Arch. Comput. Methods Eng. 28 (2021) 463–471.

[35] Poonguzhali Elangovan, Malaya Kumar Nath, Glaucoma assessment from color fundus images using convolutional neural network, Int. J. Imaging Syst. Technol. 31 (2) (2021) 955–971.

[36] Di Gong, Man Hu, Yue Yin, Tong Zhao, Tong Ding, Fan Meng, Yongli Xu, Yi Chen, Practical application of artificial intelligence technology in glaucoma diagnosis, J. Ophthalmol. 2022 (2022).

[37] Tri Huynh, Aiden Nibali, Zhen He, Semi-supervised learning for medical image classification using imbalanced training data, Comput. Methods Programs Biomed. 216 (2022) 106628.

[38] Hitendra Garg, Neeraj Gupta, Rohit Agrawal, Shivendra Shivani, Bhisham Sharma, A real time cloud-based framework for glaucoma screening using EfficientNet, Multimedia Tools Appl. (2022) 1–22.

[39] Mohammad J.M. Zedan, Mohd Asyraf Zulkifley, Ahmad Asrul Ibrahim, Asraf Mohamed Moubark, Nor Azwan Mohamed Kamari, Siti Raihanah Abdani, Automated glaucoma screening and diagnosis based on retinal fundus images using deep learning approaches: A comprehensive review, Diagnostics 13 (13) (2023) 2180.

[40] Tin Tin Khaing, Thayanee Ruennark, Pakinee Aimmanee, Stanislav Makhanov, Navapol Kanchanaranya, Glaucoma detection in mobile phone retinal images based on ADI-GVF segmentation with EM initialization, ECTI Trans. Comput. Inf. Technol. (ECTI- CIT) 15 (1) (2021) 134–149.

[41] A.P. Sunija, Varun P. Gopi, P. Palanisamy, Redundancy reduced depthwise separable convolution for glaucoma classification using OCT images, Biomed. Signal Process. Control. 71 (2022) 103192.

[42] Jiaji Wang, Shuihua Wang, Yudong Zhang, Artificial intelligence for visually impaired, Displays 77 (2023) 102391.

[43] Shahzaib Iqbal, Tariq M. Khan, Khuram Naveed, Syed S. Naqvi, Syed Junaid Nawaz, Recent trends and advances in fundus image analysis: A review, Comput. Biol. Med. (2022) 106277.

[44] Riccardo Cheloni, Jonathan Denniss, Depth-resolved variations in visibility of retinal nerve fibre bundles across the retina in enface OCT images of healthy eyes, Ophthalmic Physiol. Opt. 41 (1) (2021) 179–191.

[45] S.D. Mallanna, K. Viswanath, Pramod K.B. Rangaiah, Design and analysis of high gain planar inverted-F antenna (PIFA) for wimax and nomadic applications, in: 2021 Sixth International Conference on Wireless Communications, Signal Processing and Networking, WiSPNET, IEEE, 2021, pp. 6–10.

[46] B.P. Pradeep Kumar, Pramod K.B. Rangaiah, Robin Augustine, Enhancing medical image reclamation for chest samples using B-coefficients, DT-CWT and EPS algorithm, IEEE Access (2023).

[47] Gülcan Gencer, Kerem Gencer, Advanced retinal disease detection from OCT images using a hybrid squeeze and excitation enhanced model, PLoS One 20 (2) (2025) e0318657.

[48] Amera W. Al-funjan, Hanaa M. Al Abboodi, Najlaa Abd Hamza, Wafaa M. Salih Abedi, Alaa H. Abdullah, A lightweight deep learning-based ocular disease prediction model using squeeze-and-excitation network architecture with MobileNet feature extraction, Int. J. Intell. Eng. Syst. 17 (4) (2024).

[49] Wei Zhang, Yanxiu Li, Yixin Yu, Katherine Derouin, Yu Qin, V. Phuc Nguyen, Xiaobo Xia, Xueding Wang, Yannis M Paulus, Simultaneous photoacoustic microscopy, spectral-domain optical coherence tomography, and fluorescein microscopy multi-modality retinal imaging, Photoacoustics 20 (2020) 100194.

[50] Pramod K.B. Rangaiah, Bappaditya Mandal, Erik Avetisyan, Arvind Selvan Chezhian, Bobins Augustine, Mauricio David Perez, Robin Augustine, Preliminary analysis of burn degree using non-invasive microwave spiral resonator sensor for clinical applications, Front. Med. Technol. 4 (2022) 859498.
[51] Ines Lains, Jay C. Wang, Ying Cui, Raviv Katz, Filippos Vingopoulos, Giovanni Staurenghi, Demetrios G. Vavvas, Joan W. Miller, John B. Miller, Retinal applications of swept source optical coherence tomography (OCT) and optical coherence tomography angiography (OCTA), Prog. Retin. Eye Res. 84 (2021) 100951.
[52] Dana-Mihaela Petroşanu, Alexandru Pîrjan, Alexandru Tăbuşcă, Tracing the influence of large language models across the most impactful scientific works, Electronics 12 (24) (2023) 4957.
[53] Masoud Mohseni-Dargah, Zahra Falahati, Bahareh Dabirmanesh, Parisa Nasrollahi, Khosro Khajeh, Machine learning in surface plasmon resonance for environmental monitoring, in: Artificial Intelligence and Data Science in Environmental Sensing, Elsevier, 2022, pp. 269–298.
[54] Pramod K.B. Rangaiah, Robin Augustine, et al., Improving burn diagnosis in medical image retrieval from grafting burn samples using B-coefficients and the CLAHE algorithm, Biomed. Signal Process. Control. 99 (2025) 106814.
[55] Anita Manassakorn, Supatana Auethavekiat, Vera Sa-Ing, Sunee Chansangpetch, Kitiya Ratanawongphaibul, Nopphawan Uramphorn, Visanee Tantisevi, GlauNet: Glaucoma diagnosis for OCTA imaging using a new CNN architecture, IEEE Access 10 (2022) 95613–95622.
[56] WangMin Liao, BeiJi Zou, RongChang Zhao, YuanQiong Chen, ZhiYou He, MengJie Zhou, Clinical interpretable deep learning model for glaucoma diagnosis, IEEE J. Biomed. Heal. Inform. 24 (5) (2019) 1405–1412.
[57] Liu Li, Mai Xu, Hanruo Liu, Yang Li, Xiaofei Wang, Lai Jiang, Zulin Wang, Xiang Fan, Ningli Wang, A large-scale database and a CNN model for attention-based glaucoma detection, IEEE Trans. Med. Imaging 39 (2) (2019) 413–424.
[58] Yasmeen George, Bhavna J. Antony, Hiroshi Ishikawa, Gadi Wollstein, Joel S. Schuman, Rahil Garnavi, Attention-guided 3D-CNN framework for glaucoma detection and structural-functional association using volumetric images, IEEE J. Biomed. Heal. Inform. 24 (12) (2020) 3421–3430.
[59] Rongchang Zhao, Xuanlin Chen, Xiyao Liu, Zailiang Chen, Fan Guo, Shuo Li, Direct cup-to-disc ratio estimation for glaucoma screening via semi-supervised learning, IEEE J. Biomed. Heal. Inform. 24 (4) (2019) 1104–1113.
[60] Mir Tanvir Islam, Shafin T. Mashfu, Abrar Faisal, Sadman Chowdhury Siam, Intisar Tahmid Naheen, Riasat Khan, Deep learning-based glaucoma detection with cropped optic cup and disc and blood vessel segmentation, IEEE Access 10 (2021) 2828–2841.
[61] Ayesha Shoukat, Shahzad Akbar, Khadija Safdar, A deep learning-based automatic method for early detection of the glaucoma using fundus images, in: 2021 International Conference on Innovative Computing, ICIC, IEEE, 2021, pp. 1–6.
[62] Deepak Parashar, Dheeraj Agrawal, Improved classification of glaucoma in retinal fundus images using 2D-DWT, in: 2021 International Conference on Advances in Electrical, Computing, Communication and Sustainable Technologies, ICAECT, IEEE, 2021, pp. 1–5.
[63] Silvia Ovreiu, Irina Cristescu, Florian Balta, Elena Ovreiu, An exploratory study for glaucoma detection using densely connected neural networks, in: 2020 International Conference on E-Health and Bioengineering, EHB, IEEE, 2020, pp. 1–4.
[64] Ungsumalee Suttapakti, Supawadee Srikamdee, Janya Onpans, Multi-directional texture feature extraction for glaucoma classification from color retinal images, in: 2022 19th International Joint Conference on Computer Science and Software Engineering, JCSSE, IEEE, 2022, pp. 1–5.
[65] Pramod K.B. Rangaiah, Roger L. Karlsson, Arvind Selvan Chezhian, Laya Joseph, Bappaditya Mandal, Bobins Augustine, Maria Mani, Mauricio David Perez, Thiemo Voigt, Robin Augustine, Realization of a portable semi-shielded chamber for evaluation of fat-intrabody communication, IEEE Access 11 (2023) 72743–72755.
[66] Patthapol Kunumpol, Nichapa Lerthirunvibul, Phongphan Phienphanich, Adirek Munthuli, Kanjapat Temahivong, Visanee Tantisevi, Anita Manassakorn, Sunee Chansangpetch, Rath Itthipanichpong, Kitiya Ratanawongphaibul, Prin Rojanapongpun, Charturong Tantibundhit, GlauCUTU: Time until perceived virtual reality perimetry with humphrey field analyzer prediction-based artificial intelligence, IEEE Access 10 (2022) 36949–36962, http://dx.doi.org/10.1109/ACCESS.2022.3163845.
[67] Pramod K.B. Rangaiah, Javad Ebrahimizadeh, Jacob Velander, Roger Karlsson, Bappaditya Mandal, Mauricio D. Perez, Robin Augustine, Design of constant width branch line directional coupler for the microwave sensing application, in: 2020 14th European Conference on Antennas and Propagation, EuCAP, IEEE, 2020, pp. 1–5.
[68] Thisara Shyamalee, Dulani Meedeniya, Attention U-net for glaucoma identification using fundus image segmentation, in: 2022 International Conference on Decision Aid Sciences and Applications, DASA, IEEE, 2022, pp. 6–10.
[69] Thisara Shyamalee, Dulani Meedeniya, CNN based fundus images classification for glaucoma identification, in: 2022 2nd International Conference on Advanced Research in Computing, ICARC, IEEE, 2022, pp. 200–205.
[70] B.P. Pradeep Kumar, N.N. Srinidhi, S.L. Shiva Darshan, E. Naresh, Design of cost efficient VBIR technique using ICA and IVCA, SN Comput. Sci. 5 (5) (2024) 560.
[71] Abhilasha Joshi, K.K. Sharma, Graph deep network for optic disc and optic cup segmentation for glaucoma disease using retinal imaging, Phys. Eng. Sci. Med. 45 (3) (2022) 847–858.
[72] Jahanzaib Latif, Shanshan Tu, Chuangbai Xiao, Sadaqat Ur Rehman, Azhar Imran, Yousaf Latif, Odgnet: a deep learning model for automated optic disc localization and glaucoma classification using fundus images, SN Appl. Sci. 4 (4) (2022) 98.
[73] Pramod K.B. Rangaiah, B.P. Pradeep Kumar, Robin Augustine, Histopathology-driven prostate cancer identification: A VBIR approach with CLAHE and GLCM insights, Comput. Biol. Med. 182 (2024) 109213.
[74] Pramod K.B. Rangaiah, B.P. Pradeep Kumar, Fredrik Huss, Robin Augustine, Precision diagnosis of burn injuries using imaging and predictive modeling for clinical applications, Sci. Rep. 15 (1) (2025) 7604.
[75] Gaetano Perchiazzi, Rafael Kawati, Mariangela Pellegrini, Jasmine Liangpansakul, Roberto Colella, Paolo Bollella, Pramod Rangaiah, Annamaria Cannone, Deepthi Hulithala Venkataramana, Mauricio Perez, et al., Imitating the respiratory activity of the brain stem by using artificial neural networks: exploratory study on an animal model of lactic acidosis and proof of concept, J. Clin. Monit. Comput. (2024) 1–12.
[76] J. Ravikumar, et al., Gauging deep learning archetypal effectiveness in haematological reclamation, SN Comput. Sci. 5 (7) (2024) 1–19.
[77] B.P. Pradeep kumar, Pramod K.B. Rangaiah, Robin Augustine, Improving liver cancer diagnosis: a multifaceted approach to automated liver tumor identification in ultrasound scans, Next Research (ISSN: 3050-4759) (2025) 100465, http://dx.doi.org/10.1016/j.nexres.2025.100465.
[78] Sanjana Prasad, Arun Samimalai, S. Rashmi Rani, B.P. Pradeep Kumar, Nayana Hegde, Sufia Banu, Information security and privacy in smart cities, smart agriculture, industry 4.0, smart medicine, and smart healthcare, in: IoT Based Control Networks and Intelligent Systems: Proceedings of 3rd ICICNIS 2022, Springer, 2022, pp. 621–635.
[79] B.P. Pradeep Kumar, S.L. Shiva Darshan, E. Naresh, N.N. Srinidhi, J. Shreyas, Design of chest visual based image reclamation method using dual tree complex wavelet transform and edge preservation smoothing algorithm, SN Comput. Sci. 5 (4) (2024) 352.
[80] B.P. Pradeep Kumar, H.M. Manoj, Comparative assessment of machine learning models for predicting glucose intolerance risk, SN Comput. Sci. 5 (7) (2024) 894.
[81] Mohammad Hesam Hesamian, Wenjing Jia, Xiangjian He, Paul Kennedy, Deep learning techniques for medical image segmentation: achievements and challenges, J. Digit. Imaging 32 (2019) 582–596.
[82] Zongwei Zhou, Md Mahfuzur Rahman Siddiquee, Nima Tajbakhsh, Jianming Liang, Unet++: A nested u-net architecture for medical image segmentation, in: Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: 4th International Workshop, DLMIA 2018, and 8th International Workshop, ML-CDS 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 20, 2018, Proceedings 4, Springer, 2018, pp. 3–11.
[83] Mahavirsingh Rajpurohit, Kishor G. Sawarkar, K.B. Pramod, Kushal Tuckley, Kumarswamy, Microstrip line geometric variation consequences for linear parameters of microwave amplifiers, in: Optical and Wireless Technologies: Proceedings of OWT 2017, Springer, 2018, pp. 515–529.
[84] Le Sun, Yu Fang, Yuwen Chen, Wei Huang, Zebin Wu, Byeungwoo Jeon, Multi-structure KELM with attention fusion strategy for hyperspectral image classification, IEEE Trans. Geosci. Remote Sens. 60 (2022) 1–17.
[85] Pramod K.B. Rangaiah, H.V. Kumaraswamy, A 1-5Ghz, hybrid mic wideband LNA utilizing microstrip geometric structure variety for performance improvement, Trans. Netw. Commun. 5 (2) (2017) 15.
[86] Kaustav Sarkar, Amit Shiuly, Krishna Gopal Dhal, Revolutionizing concrete analysis: An in-depth survey of AI-powered insights with image-centric approaches on comprehensive quality control, advanced crack detection and concrete property exploration, Constr. Build. Mater. 411 (2024) 134212.
[87] Thisara Shyamalee, Dulani Meedeniya, Gilbert Lim, Mihipali Karunarathne, Automated tool support for glaucoma identification with explainability using fundus images, IEEE Access (2024).
[88] Ayesha Shoukat, Shahzad Akbar, Syed Ale Hassan, Sajid Iqbal, Abid Mehmood, Qazi Mudassar Ilyas, Automatic diagnosis of glaucoma from retinal images using deep learning approach, Diagnostics 13 (10) (2023) 1738.
[89] Kamesh Sonti, Ravindra Dhuli, A new convolution neural network model "KR-NET" for retinal fundus glaucoma classification, Optik 283 (2023) 170861.
[90] Mamta Juneja, Sarthak Thakur, Archit Uniyal, Anuj Wani, Niharika Thakur, Prashant Jindal, Deep learning-based classification network for glaucoma in retinal images, Comput. Electr. Eng. 101 (2022) 108009.
[91] K.B. Pramod, H.V. Kumaraswamy, B.P. Pradeep Kumar, C. Prathap, et al., Design and analysis of UHF bjt feedback oscillator using linear and non-linear simulation, in: 2013 International Conference on Emerging Trends in Communication, Control, Signal Processing and Computing Applications, C2SPCA, IEEE, 2013, pp. 1–6.