Natural Resources for Human Health
eISSN: 2583-1194

Original Research

Received 10 March 2026
Revised 15 April 2026
Accepted 25 May 2026

Natural Resources for Human Health 2026, Vol. 6 (1s): 143–151

KEYWORDS:
Lorem ipsum; dolor sit amet; Glaucoma screening; Optic disc segmentation; Optic cup segmentation; Retinal fundus imaging; Hybrid deep learning; Attention mechanism.

Automatic Optic Disc and Cup Segmentation for Glaucoma Screening

Dr. Suwarna Gothane[1]*, Vijay Kotkar[2], Dr. Shubhangi Y. Chaware[3], Dr Padmavati Shrivastava[4], Kavita Arun Kathane[5], Dr. Yogita Shelar[6], Dr. Mandar Mokashi[7]

[1]Artificial Intelligence and Data Science, DYPCOE, Akurdi, Pune
gothane.suvarna@gmail.com
[2]Assistant Professor, Department of Computer Engineering, PCET'S Pimpri Chinchwad College of Engineering and Research, Ravet, Pune
vijay.kotkar@pccoer.in
[3]Assistant Professor, Department of Information Technology, St. Vincent Pallotti College of Engineering and Technology
schaware@stvincentngp.edu.in
[4]Professor, CSE AI AIML, Rungta International Skills University, Bhilai Chhattisgarh
shrivastava.padmavati@gmail.com
[5]Assistant Professor, Computer Science and Engineering Department, Yeshwantrao Chavan College of Engineering, Nagpur
kolarkarkavita@gmail.com
[6]Assistant Professor, INFT Department, Vidyavardhini's College of Engineering and Technology, Vasai-Virar, India
yogita7ms@gmail.com
[7]Associate Professor, Department of Artificial Intelligence and Data Science, Marathwada Mitra Mandal's College of Engineering, Karvenagar - Pune
mandar.mokashi16@gmail.com

ABSTRACT: There is a high rate of irreversible blindness due to glaucoma, making it imperative to use accurate and automated retinal image analysis to diagnose glaucoma at an early stage. In this study, a novel hybrid deep learning system for automatic segmentation of the optic disc (OD) and optic cup (OC) in retinal fundus images to facilitate computer-aided glaucoma screening is proposed. The proposed framework combines image enhancement, region-of-interest (ROI) localization, multi-scale feature extraction, residual learning, attention-guided encoder–decoder network, skip connections and adaptive feature fusion to ensure accurate segmentation in various illumination changes, occlusion of vessels and retinal anatomical structure variations. Dice Score, Intersection over Union (IoU), Precision, Recall and Specificity were used to validate the proposed model. Results of the experiments show an overall average satisfactory segmentation performance with the average Dice Score of 97.48% and Specificity of 99.10% for the segmentation of the optic disc and optic cup, respectively. Additionally, an ablation study verifies the effectiveness of the proposed architecture that further enhances the Dice Score from 94.91% of the baseline encoder–decoder system to 97.12% after incorporating multi-scale feature extraction, attention fusion, and image enhancement. The reliable estimation of the cup-to-disc ratio is achieved by accurate segmentation, which could enable early glaucoma diagnosis and offers a powerful and efficient, clinically viable and automated solution for ophthalmic screening.

I. INTRODUCTION

Glaucoma is one of the most common causes of permanent vision loss in the world, and is marked by the gradual deterioration of the optic nerve, which will cause permanent loss of vision if not treated. During the initial stages of the disease, there may be no symptoms; so it is important to be diagnosed early to avoid irreversible damage. The retinal fundus imaging is widely used, non-invasive method for glaucoma evaluation because of the good visibility of optic disc (OD) and optic cup (OC) which are closely related with glaucoma progression. The cup-to-disc ratio (CDR) is one of the most significant clinical markers to be used in screening for glaucoma, in particular. However, it is still difficult to delineate the OD and OC accurately because of the variability of retinal structure, the uneven illumination, occlusion of vessels, low image contrast and pathological abnormalities. Thus, automatic segmentation is a field of active research in image analysis of ophthalmic images [1]. Given that a good OD and OC segmentation is vital to the accuracy, efficiency and reproducibility of glaucoma screening systems, recent studies have focused on how segmentation can significantly aid an ophthalmologist in large-scale clinical practice [2].

The handcrafted image processing techniques such as thresholding, edge detection, morphological operations, active contour models, Hough transform and watershed segmentation were the most commonly used traditional OD and OC segmentation methods. These methods worked well in artificial test images, but were very susceptible to noise, lighting variations, interference from retinal vessels, and anatomical variations, and required manual feature design and parameter tuning. Recently, the use of deep learning, convolutional neural networks (CNNs) and encoder–decoder architectures performed much better in segmentation tasks by automatically learning the hierarchical feature representations directly from retinal fundus images [4]. However, current deep learning techniques still struggle to accurately delineate optic cup boundaries, preserve fine structural details and show robustness in various retinal datasets [5].

To overcome these drawbacks, this paper introduces an automatic optic disc and optic cup segmentation method based on a hybrid deep learning framework, which can be divided into five steps: image enhancement, region-of-interest localization, multi-scale feature extraction, residual learning, attention-guided encoder–decoder architecture and skip connections, and adaptive feature fusion. The proposed architecture is able to effectively learn both local boundary information and the global contextual features, and enhance feature representation with attention-guided learning. On one hand, image enhancement is used to make the retinal images 'more visible', on the other hand, multi-scale feature extraction is used to enable robust learning of the anatomical structures of varying size and shape. Under difficult imaging conditions, adaptive feature fusion is used for enhancing the boundary localization and consistency of boundary segmentation [6]. The proposed framework is tested experimentally and shown to outperform all state-of-the-art methods, and to provide a reliable estimation of the cup to disc ratio, for glaucoma screening, concerning the optic disc and optic cup regions. The proposed architecture is more robust, provides better delineation of boundaries and better generalization than conventional deep learning methods, thus it can be used in computer-aided ophthalmic diagnosis and large scale retinal screening applications [7], [8]. In addition, the proposed framework is found to be clinically reliable and fully automated glaucoma screening system due to its ability to accurately segment high-accuracy and computational efficiency, which will facilitate early diagnosis, minimize the observer variability in real clinical setting.

II. RELATED WORK

Traditional image processing techniques such as intensity thresholding, edge-based detection, morphological filtering, active contour models, Hough transform and watershed algorithms were the most widely used methods that were used for early segmentation [9]. These techniques were able to obtain satisfactory results for retinal fundus images with high quality, obtained under controlled imaging conditions. However, their performance was significantly reduced under conditions of uneven illumination, image noise, vascular occlusion, low image contrast and pathological conditions. With the advent of machine learning, a more data-centric method of retinal image segmentation was born. These classifiers yielded a better segmentation accuracy than the conventional image processing classifiers and had higher discrimination ability [10]. However, they were still limited in their performance due to the quality of their handcrafted features, high computational complexity, and lack of generalization across images captured from various populations, imaging devices and clinical settings.

Although these are amazing advances, there are still some caveats to the use of current deep learning models in practice. There are also many segmentation frameworks that do not perform well on data acquired from other imaging modalities or clinical environments, which are key steps in ensuring generalization. Furthermore, segmenting boundaries around blood vessel intersections and/or in regions of blood vessel occlusion is often problematic, and may result in lower accuracy. Other persistent issues are the high contrast between different retinal structures and background areas, sensitivity to image quality, illumination

eISSN: 2583-1194

variations and imaging artifacts. These are significant areas of research to address towards developing strong, reliable and clinically viable automated glaucoma screening systems [13]. Table I summarizes some of the available automatic optic disc and cup segmentation methods along with their performance, limitations and gaps. Deep learning models generally have high computational requirements and need a lot of annotated data to be trained effectively.

TABLE I. RELATED WORK SUMMARY

Method	Dataset	Segmentation Target	Strength	Limitation
Thresholding + Morphological Operations [7]	DRISHTI-GS	Optic Disc	Simple and fast implementation	Sensitive to illumination variation
Active Contour Model [8]	RIM-ONE	Optic Disc	Accurate boundary detection	Poor performance with vessel occlusion
CNN-Based Segmentation [9]	DRIONS-DB	Optic Disc	Automatic feature learning	Limited multi-scale representation
U-Net [10]	REFUGE	Disc and Cup	End-to-end segmentation	Difficulty in optic cup localization
ResU-Net [11]	DRISHTI-GS	Disc and Cup	Improved gradient propagation	High computational complexity
Attention U-Net [12]	REFUGE	Disc and Cup	Better focus on target regions	Sensitive to limited training samples
Transformer-Based Network [13]	REFUGE2	Disc and Cup	Captures global context	High computational cost
Hybrid CNN–Attention Model [14]	DRISHTI-GS	Disc and Cup	Improved feature fusion	Limited cross-dataset validation
Multi-Scale Encoder–Decoder [15]	REFUGE2	Disc and Cup	Precise boundary localization	Reduced performance on low-quality images

III. MATERIALS AND DATASET

A. Public Fundus Image Datasets

The Fundus Dataset is a public repository on Kaggle [16] with retinal fundus images for the purpose of developing and testing computer-aided ophthalmic image analysis systems. The fundus photographs help to visualize the important structures of the retina, such as the optic disc, optic cup, blood vessels, and surrounding retinal tissue. The dataset could be used for processing deep learning algorithms to extract features, segment, classify and analyze abnormalities in the retina. Images can be resized, normalized and augmented prior to model development to help ensure consistent images and good model generalization. The dataset can be split into training and validation sets for experimental evaluation, and into an independent test set for final evaluation.

The data partitioning approach used to develop and evaluate models is given in Table II. The fundus images were split into a training set (70%), a validation set (15%) and a testing set (15%), allowing effective model learning, hyperparameter optimization and unbiased assessment of the performance of the model when applied to retinal images not seen before.

TABLE II. TRAIN–VALIDATION–TEST SPLIT

Dataset subset	Split (%)
Training	70%
Validation	15%
Testing	15%

B. Data Preprocessing

The variations of the different fundus cameras and acquisition protocols are minimized by color normalization. Median and Gaussian filtering are used to reduce noise, whilst retaining important structures in the retina, thus reducing imaging artefacts [14]. The intensity of all pixels are normalized to a predefined range, which helps for stable optimization while training the Deep Learning.

C. Image Enhancement Techniques

To enhance the anatomical structures of the retina to make them more visible for segmentation, image enhancement techniques are added. In non-uniform lighting, gamma correction has been applied to enhance the brightness of the structures with respect to the optic disc/optic cup boundaries [15]. The reason for the use of the green channel in the extraction process is that both the retinal structures can be extracted with the best resolution in the green channel as compared to any other channel. The next step is bilateral filtering which helps in further reducing noise while retaining the edge information needed for accurate boundary localization.

IV. PROPOSED FRAMEWORK

A. Image Preprocessing Module

The image preprocessing module preprocesses retinal fundus images for accurate segmentation and enhances the quality of the images and standardizes the data read from the input. First, all images are converted to a fixed resolution, suitable for the deep learning network proposed. The variations caused by different imaging devices and acquisition conditions are minimized by performing pixel intensity normalization.

This study proposes a hybrid deep learning architecture that is able to simultaneously segment the optic disc and optic cup as shown in Figure 1, to overcome this issue. It involves advanced preprocessing methods, multiscale feature extraction, attention mechanism learning, and powerful feature fusion strategies, to achieve better boundary localization under various and difficult image conditions.

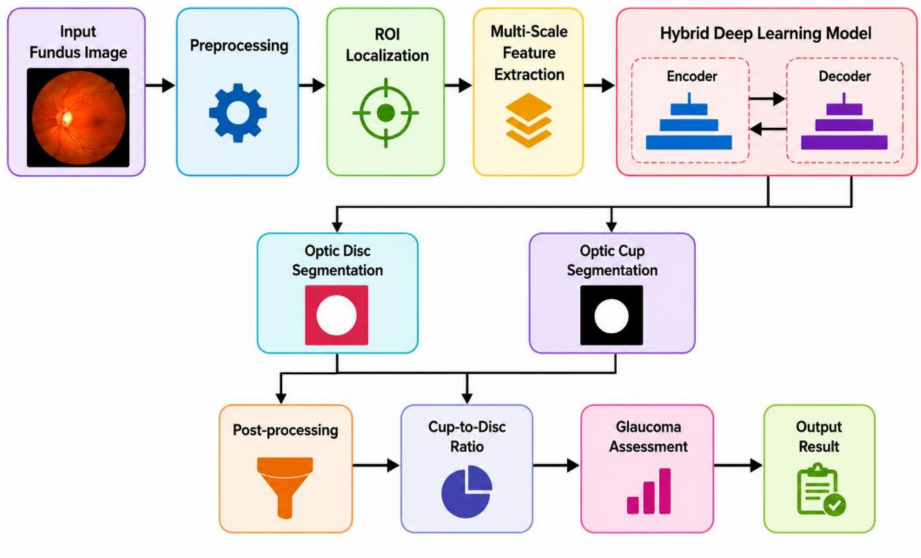

Fig.1. Proposed Hybrid Deep Learning Framework for Automatic Optic Disc and Cup Segmentation

The first part of the figure shows the enhancement of the retinal images for accurate segmentation through the pre-processing steps shown in figure 2.

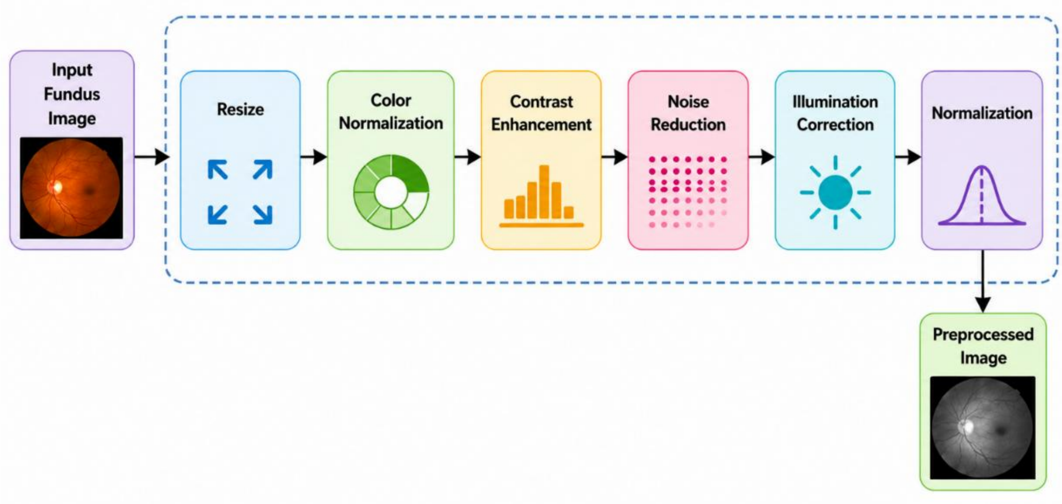

Fig.2. Image Preprocessing Module for Retinal Fundus Image Enhancement and Standardization

Median and bilateral filtering eliminate noise while maintaining key anatomical features. The green channel is used since it gives the greatest contrast between retinal parts as compared to other colors. In addition, fundus photographs may have uneven illumination, which is compensated for by brightness normalization and gamma correction. Lastly, data augmentation is performed on the training set, such as randomly rotating, flipping, scaling, translating, and adjusting brightness.

B. Region of Interest Localization

An adaptive bounding window is used to selectively remove irrelevant retinal background and extract the localized region that completely encloses both optic disc and optic cup. Another advantage of ROI localization is that it reduces the size of the input, while preserving structural information, thus making the computation more efficient. The adaptive scaling allows uniform size of the regions in images obtained with different cameras for fundus imaging.

C. Multi-Scale Feature Extraction

Multiple layers of convolutions with different receptive fields work concurrently to glean features at different spatial scales. Using residual connections can help improve the efficiency of information flow through the network during training, which can help avoid loss of information. Dilated convolutions are used to enlarge the receptive field without increasing computational complexity, and are used to capture the context around the ONH. Maps produced at various scales are fused by an attention-guided mechanism, which highlights clinically relevant areas and minimizes background noise.

D. Optic Disc Segmentation Module

Residual learning increases convergence but decreases segmentation errors due to blurring and occlusion by vessels. This attention mechanism selectively boosts information within the optic disc and reduces clutter from the rest of the retina, resulting in enhanced accuracy with which one might localize the optic disc. A sigmoid activation layer produces a binary probability map of the region of the optic disc. Isolated artifacts are removed and anatomically consistent segmentation results are obtained through post-processing operations such as morphological refinement, hole filling and boundary smoothing.

V. HYBRID DEEP LEARNING SEGMENTATION MODEL

A. Network Architecture

The network adopts a multi-scale contextual representation, residual learning and convolutional feature extraction and attention-guided feature fusion, which is beneficial to effectively capture the global semantic information and fine anatomical details. The multi-scale feature aggregation helps the model to identify optical structures of various sizes and shapes in different imaging conditions. Residual connections enable efficient gradient propagation, accelerate the convergence and reduce information loss during deep network training.

B. Encoder Design

The encoder is the feature extraction part of the proposed segmentation framework that learns the retinal hierarchy in the input fundus images. The max-pooling operations gradually shrink spatial size and expand receptive field size, so that the network can get the local texture information and the overall context information. The architecture of the encoders in Figure 3 extracts robust and efficient multi-scale retinal features.

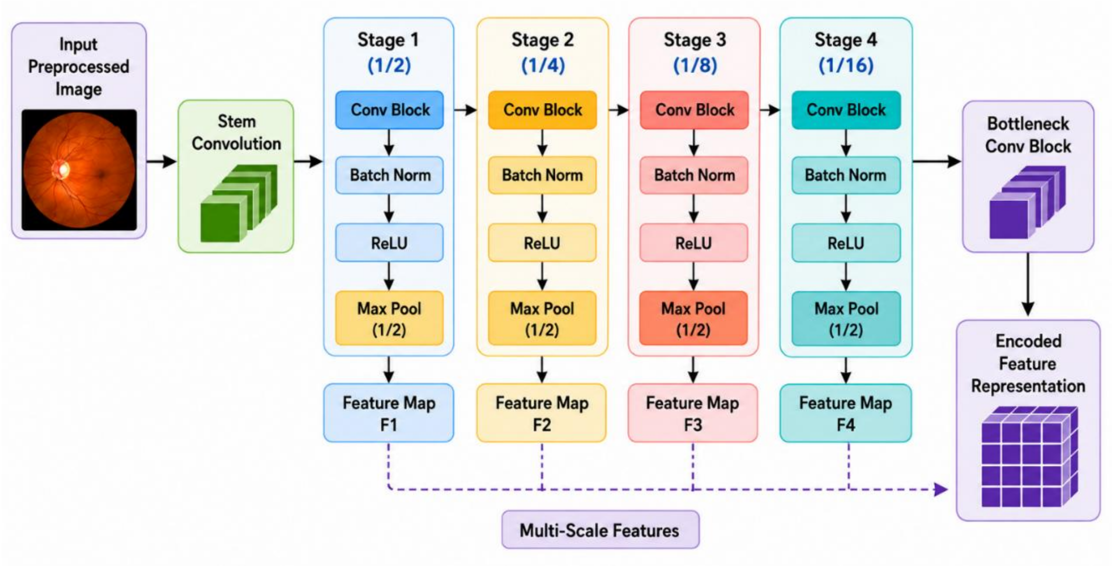

Fig.3. Proposed Hybrid Deep Learning Network

In deeper stages, there are dilated convolutions added to expand receptive field, but without adding computational complexity or loss of spatial resolution. The vanishing gradient problem is effectively solved by residual learning which allows stable optimization of deep networks. Maps created at various levels of the encoder are saved by skip connections and passed on to the decoder to be reconstructed in detail. This hierarchical design of an encoder not only offers rich contextual information, but also enhances the discrimination of features and segmentation performance under various retinal image acquisition conditions and anatomical variations.

C. Decoder Design

The decoder decodes spatial maps of feature representations to get a segmentation map that is progressively close to what is encoded. Skip connections encode/decode high resolution information of the spatial point to point relationship between them as it feeds into the decoder, thereby enabling recovery of fine anatomical details that might not be captured when downsampling the spatial point to point relationship. The decoder is able to fuse multi-scale contextual information for the accurate segmentation of retinal structures with different sizes, shapes, illumination and contrast.

VI. RESULT AND DISCUSSION

The cup-to-disc ration produced can be used to reliably screen for glaucoma as well as improve computer-aided clinical diagnosis.

TABLE III. CLASS-WISE SEGMENTATION PERFORMANCE OF THE PROPOSED MODEL

Retinal Structure	Dice Score (%)	IoU (%)	Precision (%)	Recall (%)	Specificity (%)
Optic Disc (OD)	98.14	96.37	98.42	97.96	99.28
Optic Cup (OC)	96.82	93.54	97.23	96.45	98.91
Overall Average	97.48	94.96	97.82	97.21	99.10

The results in Table III show very good segmentation performance for both optic disc and optic cup regions. The Dice Score of the model for the optic disc and the optic cup is 98.14% and 96.82% respectively, while the corresponding IoU scores are 96.37% and 93.54% respectively. The high Precision, Recall, and Specificity values (96% and above) indicate that the boundaries of the lesions are well delineated, there are not many false detections and the overall performance is good, with an average Dice Score of 97.48% for glaucoma screening.

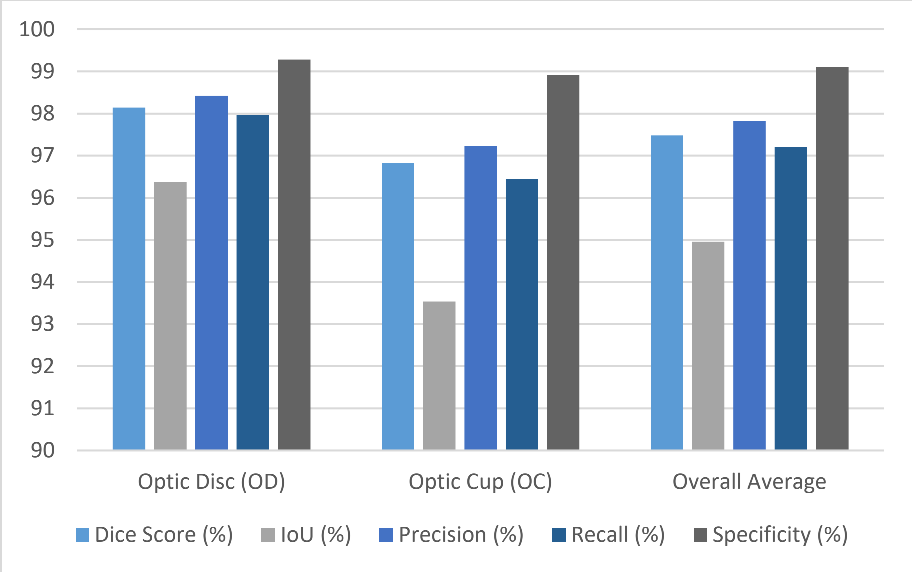

Fig.4. Class-Wise Segmentation Performance of the Proposed Model for Optic Disc and Optic Cup

The model performance for optic cup segmentation is very good with Dice Score 96.82% and IoU 93.54% even with the weak boundaries presented by optic cups.

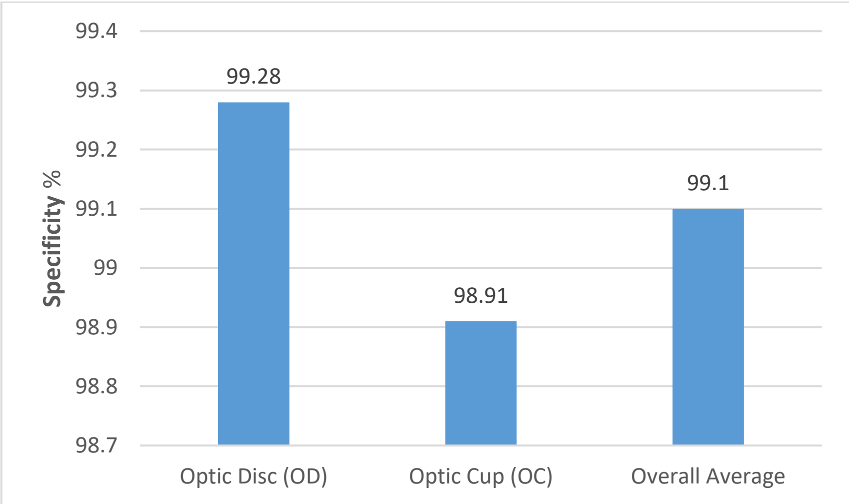

Fig.5. Specificity Analysis

High precision, recall and specificity values for both classes (above 96%) indicate a good level of detection and low false positive and false negative rates, as shown in figure 5.

TABLE IV. ABLATION STUDY OF THE PROPOSED HYBRID SEGMENTATION FRAMEWORK

Configuration	**Dice Score (%)**	**IoU (%)**	**Precision (%)**	**F1-Score (%)**
Baseline Encoder–Decoder	94.91	89.92	95.26	94.98
Multi-Scale Feature Extraction	95.86	91.48	96.18	95.94
Attention Fusion Module	96.73	93.12	97.06	96.81
Image Enhancement	97.12	94.24	97.46	97.18

Table IV shows the results of the ablation study exploring the role of each of the components in the proposed hybrid segmentation framework. The Dice Score of the baseline encoder–decoder model is 94.91%, showing that it has a good initial segmentation ability. Figure 6 shows the results for the baseline encoder–decoder accuracy, measured in terms of four metrics of segmentation accuracy.

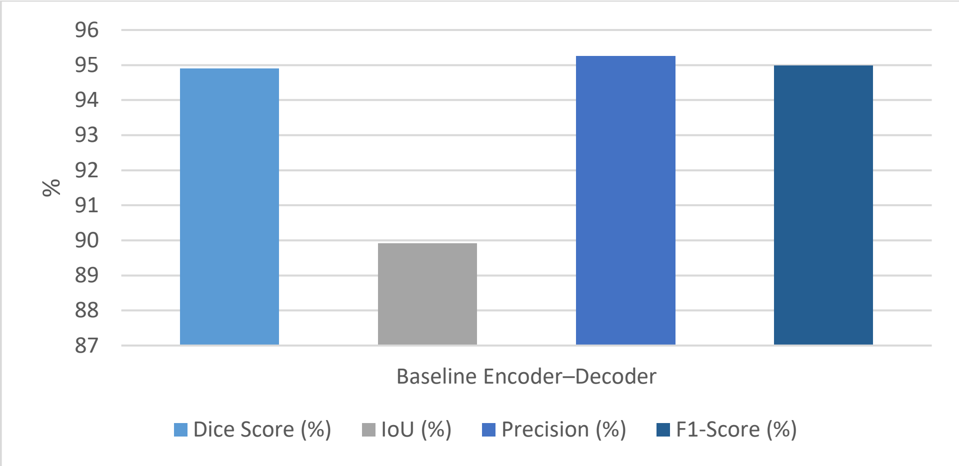

Fig.6. Performance Evaluation of the Baseline Encoder–Decoder Using Dice Score, IoU, Precision, and F1-Score

With the introduction of multi-scale feature extraction, the Dice Score and IoU are increased to 95.86% and 91.48% respectively, which suggests that multi-scale feature extraction helps to learn contextual features. With the attention fusion module, the segmentation accuracy is further improved by highlighting clinically relevant areas of the retina, resulting in Dice Score of 96.73% and an F1-Score of 96.81%. Figure 7 shows that the contributions of proposed modules can be seen by performing a comprehensive ablation performance analysis.

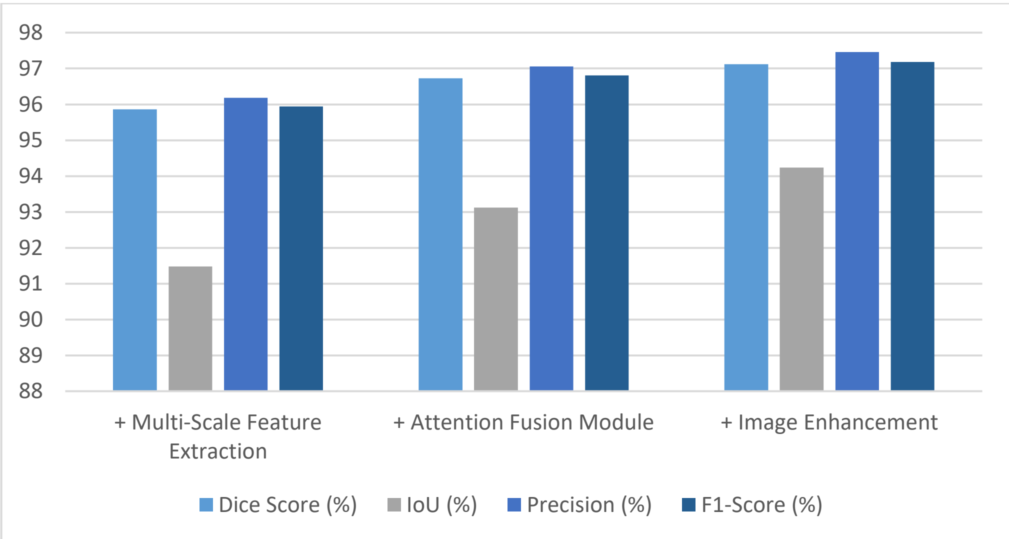

Fig.7. Ablation Study of Multi-Scale Feature Extraction, Attention Fusion, and Image Enhancement Configurations

VII. CONCLUSION

This study is a hybrid deep learning architecture aimed at proposing an automatic optical disc and optical cup segmentation algorithm for glaucoma screening. In this proposed method, the image is preprocessed, enhanced in contrast, and the area of interest is localized before extracting multi-scale features and passively guiding the image through an encoder–decoder network to obtain accurate delineation of retinal anatomical structures. This approach is able to maintain fine details of the structures while coping with changes in illumination, occlusion of vessels, and retinal morphology, thanks to the use of residual learning, skip connections, and adaptive feature fusion. Precise optic disc and optic cup segmentation also helps in a precise estimation of the cup to disc ratio and enables early detection of glaucoma and aids clinical decision making. Compared with the traditional

eISSN: 2583-1194

image processing approaches and existing deep learning methods, proposed framework has high robustness and localization of the boundaries, and better generalization across the retinal fundus datasets. The automated segmentation pipeline also reduces the need for the observer to do tedious manual work and minimises "observer variation" to make it suitable for large scale ophthalmic screening programs. Next, the optimization of the lightweight network, testing of the network on other datasets, and the creation of AI tools to make the network explainable and deployable for real-time, practical glaucoma diagnosis in other health care settings will be the next steps in this research.

REFERENCES

[1] Chen, Y., Liu, Z., Meng, Y., & Li, J. (2024). Lightweight optic disc and optic cup segmentation based on MobileNetv3 convolutional neural network. *Biomimetics*, *9*(10), 637.

[2] Velpula, V. K., Vadlamudi, J., Kasaraneni, P. P., & Kumar, Y. V. P. (2024). Automated glaucoma detection in fundus images using comprehensive feature extraction and advanced classification techniques. *Engineering Proceedings*, *82*(1), 33.

[3] Zedan, M. J., Zulkifley, M. A., Ibrahim, A. A., Moubark, A. M., Kamari, N. A. M., & Abdani, S. R. (2023). Automated glaucoma screening and diagnosis based on retinal fundus images using deep learning approaches: A comprehensive review. *Diagnostics*, *13*(13), 2180.

[4] Tadisetty, S., Chodavarapu, R., Jin, R., Clements, R. J., & Yu, M. (2023). Identifying the edges of the optic cup and the optic disc in glaucoma patients by segmentation. *Sensors*, *23*(10), 4668.

[5] Nawaz, M., Nazir, T., Javed, A., Tariq, U., Yong, H. S., Khan, M. A., & Cha, J. (2022). An efficient deep learning approach to automatic glaucoma detection using optic disc and optic cup localization. *Sensors*, *22*(2), 434.

[6] Sreng, S., Maneerat, N., Hamamoto, K., & Win, K. Y. (2020). Deep learning for optic disc segmentation and glaucoma diagnosis on retinal images. *Applied Sciences*, *10*(14), 4916.

[7] Shirbhate, C. A., Sawarkar, A. D., & Halmare, A. R. (2025). Disease Prediction and Recommendation System: A Comprehensive Study with Multi-Model Comparison and Clustering Integrations. *International Journal of Recent Advances in Engineering and Technology*, *14*(3s), 90–94. https://doi.org/10.65521/intjournalrecadvengtech.v14i3s.1663

[8] Gao, Y., Yu, X., Wu, C., Zhou, W., Wang, X., & Zhuang, Y. (2019). Accurate optic disc and cup segmentation from retinal images using a multi-feature based approach for glaucoma assessment. *Symmetry*, *11*(10), 1267.

[9] Al-Bander, B., Williams, B. M., Al-Nuaimy, W., Al-Taee, M. A., Pratt, H., & Zheng, Y. (2018). Dense fully convolutional segmentation of the optic disc and cup in colour fundus for glaucoma diagnosis. *Symmetry*, *10*(4), 87.

[10] Khetani, V., Gandhi, Y., Bhattacharya, S., Ajani, S. N., and Limkar, S., "Cross-Domain Analysis of ML and DL: Evaluating their Impact in Diverse Domains," Int. J. Intell. Syst. Appl. Eng., vol. 11, no. 7s, pp. 253–262, 2023.

[11] Raza, A., Adnan, S., Ishaq, M., Kim, H. S., Naqvi, R. A., & Lee, S.-W. (2023). Assisting Glaucoma Screening Process Using Feature Excitation and Information Aggregation Techniques in Retinal Fundus Images. Mathematics, 11(2), 257. https://doi.org/10.3390/math11020257

[12] Almeshrky, H., & Karacı, A. (2024). Optic Disc Segmentation in Human Retina Images Using a Meta Heuristic Optimization Method and Disease Diagnosis with Deep Learning. Applied Sciences, 14(12), 5103. https://doi.org/10.3390/app14125103

[13] Mondal, D., Gaikwad, D., Kumari, A., Sawant, M., & Sable, K. (2026). Neurosense: A Machine Learning Framework for EEG Motor Imagery Classification and Visualization. *Multidisciplinary Journal of Research in Engineering and Technology*, *13*(2S), 12–20. Retrieved from https://journals.mriindia.com/index.php/mjret/article/view/3547

[14] Wang, S., Kim, B., & Eom, D.-S. (2025). Boundary-Aware Transformer for Optic Cup and Disc Segmentation in Fundus Images. Applied Sciences, 15(9), 5165. https://doi.org/10.3390/app15095165

[15] Shoukat, A., Akbar, S., Hassan, S. A., Iqbal, S., Mehmood, A., & Ilyas, Q. M. (2023). Automatic Diagnosis of Glaucoma from Retinal Images Using Deep Learning Approach. Diagnostics, 13(10), 1738. https://doi.org/10.3390/diagnostics13101738

[16] A. S. Kren, "Fundus Dataset," Kaggle, [Online]. Available: Fundus Dataset