Journal of Innovative Image Processing (ISSN: 2582-4252)
www.irojournals.com/iroiip/

Automatic Computer-Aided Diagnosis Tool for Glaucoma Detection Using U-Net and CNN on Retinal Fundus Images

Vaibhav Yadav[1], Barnali Dey[2], Saumya Das[3*], Udayan Baruah[4], Hashinur Islam[5], Suman Das[6]

[1,2]Department of Information Technology, Sikkim Manipal Institute of Technology, Sikkim Manipal University, Majitar, Sikkim, India.
[3]Department of Computer Science and Engineering – Artificial intelligence, Brainware University, Barasat, West Bengal, India.
[4]Birangana Sati Sadhani Rajyik Vishwavidyalaya (A state university under Government of Assam), Golaghat, Assam, India.
[5]Department of Electronics and Communication Engineering, HKBK College of Engineering, Bengaluru, Karnataka, India.
[6]Department of Computer Science and Engineering, Brainware University, Barasat, West Bengal, India.

E-mail: [1]vyadav329@gmail.com, [2]barnali.d@smit.smu.edu.in, [3*]saumya.das.1976@gmail.com, [4]udayanbaruah@yahoo.co.in, [5]hashinur0001@gmail.com, [6]das.suman194@gmail.com

Abstract

Glaucoma is a progressive optic neuropathy and a leading cause of irreversible blindness worldwide, where vision loss can be prevented only through timely diagnosis and intervention. Therefore, there is a requirement for the development of an automatic screening system to detect glaucoma at its preliminary stages. In this work, a fully automated computer-aided diagnosis (CAD) framework is proposed for early glaucoma detection using color retinal fundus images. The proposed method integrates image validation, region extraction, data balancing, data preprocessing, structural segmentation, and deep learning-based classification. The optic disc and optic cup are then segmented with a U-Net–based architecture, which allows calculation of important anatomic measures such as the vertical cup-to-disc ratio and neuroretinal rim parameters. These handcrafted features along with deep feature representations are used to train a custom convolutional neural network classifier to discriminate between glaucomatous and healthy eyes. We test the proposed system on publicly available benchmark datasets and show that it has strong robustness and generalization capacity. The achieved accuracy, sensitivity, specificity and AUC are 88.38%, 84.95%,81.43% and 93.24% respectively, in the validation dataset. The whole CAD system is applied with an easy-to-use clinical interface implemented with Streamlit to realize real-time inference and be easily integrated into the screening workflow. The proposed method is a cost-effective and reliable technique for glaucoma screening and has significant potential for mass-screening and tele-ophthalmology applications, especially in resource-limited environments.

Keywords: Glaucoma, Computer-Aided Diagnosis, Fundus Images, U-Net, CNN, Optic Disc and Cup Segmentation.

*** Corresponding Author**

DOI: https://doi.org/10.36548/jiip.2026.1.008
Received: 08.12.2025, received in revised form: 05.01.2026, accepted: 19.01.2026, published: 30.01.2026

1. Introduction

Glaucoma is a permanent, slow-moving, and progressive eye disease that eventually harms the optic nerve, most commonly due to high intraocular pressure. Due to its silent development, early detection is the only way to save patients, as most do not discover their condition until very advanced stages of vision loss. The number of glaucoma patients worldwide was anticipated to be 79.6 million in 2020 and is projected to reach 111.8 million in 2040. Moreover, Asia accounts for 87% of angle closure glaucoma (ACG) cases and 47% of the total number of glaucoma cases. People over 60 years of age are the most at risk, and glaucoma is still the second major cause of blindness globally [1]. Early identification is therefore essential to prevent irreversible structural and functional damage. Fundus images, images of the interior surface of the eye, play a vital role in glaucoma screening. Traditionally, ophthalmologists analyze these images by inspecting the optic disc, optic cup, and cup-to-disc ratio (CDR) to identify subtle retinal changes.

Nevertheless, manual evaluation takes a lot of time, is influenced by the evaluator's point of view, and is susceptible to variations in different cases. In a retinal fundus image, the optic disc (OD) is the place where the main blood vessels and nerve fibers enter, and the optic cup (OC), the central bright area, is the part that is structurally most changed due to glaucomatous " cupping." As shown in Figure 1, the enlargement of the optic cup increases the CDR, making it a widely used diagnostic biomarker [2]. Various deep learning techniques have been extensively applied to optic cup (OC) and optic disc (OD) segmentation [3]. Additionally, deep learning approaches—particularly convolutional neural networks (CNNs)—have been rigorously employed for glaucoma detection [4].

This research demonstrates an automatic computer-aided system for glaucoma detection from retinal fundus images. This work is novel in its fruitful integration of many techniques such as image validation, region extraction, data balancing, data preprocessing, structural segmentation, and deep learning-based classification. While existing research mainly focused on either segmentation or classification alone, this method proposes:

1. LeNet-based fundus image validation to automatically exclude non-fundus or low-quality images before analysis.
2. Brightest-spot-guided ROI extraction to localize the optic disc robustly under illumination variations.
3. Feature-space ADASYN balancing applied only to training data to handle class imbalance without altering image anatomy.
4. Contrast Limited Adaptive Histogram Equalization (CLAHE) applied as a pre-processing technique to improve the image quality.
5. U-Net-based optic disc and cup segmentation for clinically meaningful CDR computation.
6. CNN-based glaucoma classification fused with CDR-based decision using majority voting, improving reliability and reducing false predictions.

This end-to-end, code-validated fusion framework is the key research contribution and differentiates the proposed work from prior glaucoma CAD systems. The selection of

techniques was guided by dataset size, class imbalance, the requirement for pixel-level segmentation, clinical relevance, interpretability, and computational efficiency.

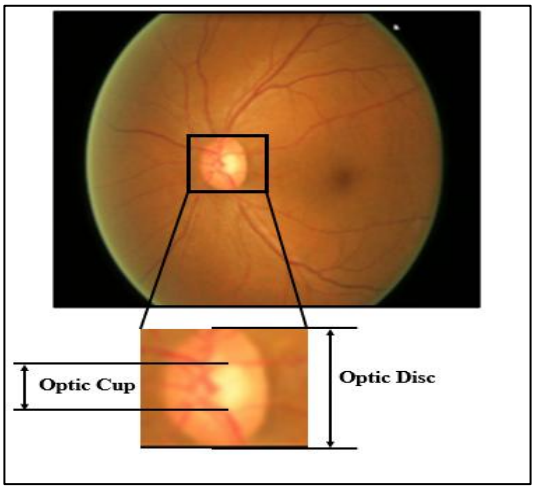

Figure 1. Retinal Fundus Image

2. Related Work

Glaucoma identification from retinal fundus pictures is a field of research where broad and deep learning methods, both classical and deep learning–based, have been used for this purpose. Typically, image processing methods extract structural features such as vascular density, retinal texture, and the cup-to-disc ratio (CDR), which is the most important biomarker of optic nerve damage. Although these handcrafted methods can be effective, their performance depends heavily on accurate feature selection and segmentation, making them sensitive to image quality and dataset variability. Convolutional Neural Networks (CNNs) have gained widespread application in medical image processing by offering the ability to learn discriminative features by themselves on large visual datasets, as machine learning advances. Hierarchy enables them to record subtle changes in the retina, which makes it possible to distinguish clearly between glaucomatous and non-glaucomatous images [5] U-Net has now established itself as a leading architecture in segmentation due to its encoder-decoder architecture and ability to localize anatomical features, such as the optic disc and cup, despite having limited training samples. Although deep models have achieved success, issues of interpretability and integrating them into actual clinical workflows remain [6]. To resolve the usability issue, several studies suggest using a combined U-Net and CNN pipeline to analyze glaucoma [7]. The goal of such hybrid systems is to utilize segmentation to gain a structural understanding and classification to achieve better diagnostic accuracy. Graphical user interfaces frequently promote these techniques to improve interface accessibility among clinicians [8]. Classical methods of optical disc localization are also widely researched. They can be square-region extraction cantered around an estimated cup centre, intensity-based OD localization, and geometric operations to optimize ROI positioning [9]. OD boundary extraction has been performed using edge-based methods, including the Laplace, Sobel, and Canny operators. Canny typically provides better localization accuracy. Variance-based thresholding, low-pass filtering, and Hough transforms are also employed in other works for template-based OD segmentation. In most cases, elliptical templates are found to be the most effective [10]. Other methods use curvature-texture, histogram equalization, morphological, and P-tile thresholding to identify OD and OC regions, and then estimate CDR to detect

glaucoma [11]. Contour-based morphological and ROI-based component labeling have been employed in segmentation-based OD/OC detection, utilizing the red and green channels of fundus images to extract OD and OC, respectively. The works indicate the huge diversity of methodology in the field. The development of deep learning-based classification has been achieved through the introduction of numerous CNN architectures. A comparison of various CNNs supports the necessity of standard databases to ensure fair benchmarking [12]. One of the significant difficulties is determining the best structures, as the performance of CNN is affected by the depth of the network, hyperparameters, and the training method. As a result, neuroevolutionary-based methods have been introduced to automate the process of designing CNN architectures, and they have been cited as a source of inspiration in the field of medical imaging [13].

Recent papers use U-Net and Fully Convolutional Networks (FCNs) to perform OD/OC segmentation and combine them with CNN-based classification. Numerous works examine CNN variants trained with a small number of epochs [14], CNN features combined with SVMs, or hybrid systems that integrate CNNs, recurrent CNNs, and FCNs to achieve better performance [15]. Deep residual networks are employed by other researchers and have proven to be better than classical models like SVM and Random Forest [16]. Other papers derive features of deeper CNNs (e.g., 18-layer networks) with SoftMax classifiers [17]. CNNs vary in architecture and that datasets differ also encourages the use of standardized datasets in the benchmarking of systems, a common practice in comparative studies of glaucoma detection systems.

3. Data

This research utilized the retinal fundus dataset that is publicly available on Kaggle [26]. The dataset contains color fundus images divided into two folders: Glaucoma_Positive and Glaucoma_Negative. . The image paths and corresponding class labels were obtained by running the code that used Python's glob and os libraries. A total of 650 images were utilized, comprising 482 non-glaucoma images (class 0) and 162 glaucoma images (class 1). Because this distribution exhibited substantial class imbalance - a common issue in medical imaging - the minority glaucoma class was augmented using the Adaptive Synthetic Sampling (ADASYN) technique [18]. ADASYN utilizes synthetic samples where the minority class is not well represented, particularly in the decision boundary, thereby enhancing the model's learning capacity on glaucoma-specific patterns. The total number of glaucoma class images post-augmentation was 506 (162 natural and 344 synthetic), and the control group consisted of 482 images, totalling 988 images for training. A separate group of 198 original non-synthetic images was set aside for testing and validation to provide a robust performance assessment. Such images likely represent clinically validated samples, enabling a realistic evaluation of diagnostic accuracy, sensitivity, and specificity. All images were also normalized to 256 × 256 pixels, as it is customary in the analysis of retinal images [19]. The combined use of original and ADASYN-enhanced samples provided a balanced and robust training environment for the proposed glaucoma classification model.

4. Methodology

The designed CAD pipeline includes the six significant procedural elements detailed in Sections 4.1, 4.2, 4.3, 4.4, 4.5 and 4.6. Image preprocessing is the first step in the methodology,

which essentially comprises grayscale conversion, contrast enhancement, and Gaussian-based denoising, followed by Region of Interest (ROI) extraction. U-Net segmentation and feature extraction were originally part of earlier experiments, but in the final pipeline, ROI-based structural representation and CNN-based classification were used for glaucoma prediction. The generalized workflow of the CAD system is shown in Figure 2.

4.1 Input Image Validation

To ensure that an input image undergoes processing only it is of the retinal fundus, an input image validation module is used before detection of glaucoma. The CNN classifier, constructed on LeNet is used to differentiate fundus and non fundus images. This step will avoid unwanted or noisy pictures from entering the pipeline. This validation experiment was performed on a dataset of 198 images. Figure 3 presents the LeNet architecture with 256 × 256 pixel grayscale images as input. Its choice was encouraged by (a) a rapid training speed, (b) the absence of GPU use, and (c) a small architecture, which is suitable for binary classification. LeNet is made up of two convolutional blocks, active blocks, pooling blocks, and fully connected layers (FC), which are succeeded by a SoftMax classifier. The model was trained on resized fundus images of 256×256 using stochastic gradient descent with a momentum of 0.95, a fixed learning rate of 0.0004, a batch size of 16, for 150 epochs, with an 80/20 train–test split. It employs ReLU activation and categorical cross-entropy loss and is optimized by the Adam optimizer. The final prediction is associated with the most probable class. LeNet was developed to identify digits but is also successful in learning fundus-specific patterns of space using convolutional layers and fully connected dense blocks to identify valid images for further processing.

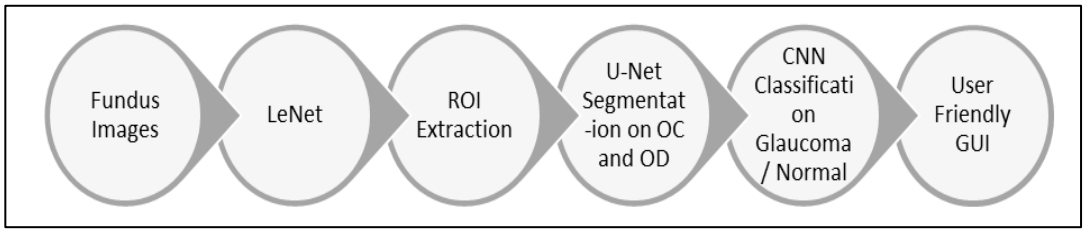

Figure 2. Model Development Flow

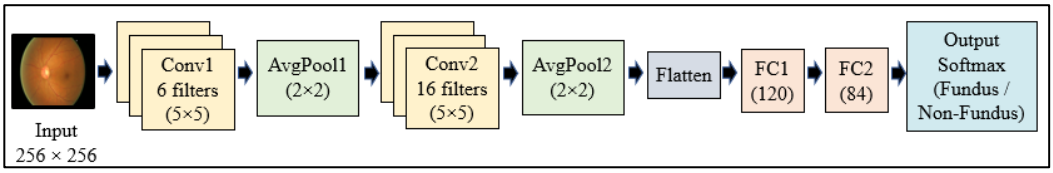

Figure 3. LeNet for Softmax Classification

4.2 ROI Extraction

The area around the optic disc is called the Region of Interest (ROI) and it is of utmost importance in the process of glaucoma assessment [20]. As Figure 4 demonstrates, ROI extraction takes less time to calculate than a regular extraction and is more precise, as it creates a model out of the most informative part of the retina. One algorithm that calculates ROI is a brightest-spot algorithm based on the observation that in a fundus image, the optic cup will tend to contain the brightest pixels. Owing to the reliance on the relative-intensity contrast method and the illumination normalization that introduce it, brightest-spot-based ROI extraction is independent of changing lighting environments. This input is then converted to grayscale and

the pixel with the highest intensity is identified. An area around this pixel is then clipped to a certain radius. The conventional technique is however vulnerable to light noise and brightness changes. Gaussian blurring is carried out to strengthen the location of bright-spots. The high spatial frequency noise is smoothed with a trial-and-error-selected Gaussian filter kernel size (65 × 65) and sigma (0). This significantly enhances ROI localization. ROI extraction accuracy was validated using localization success and overlap consistency with ground-truth optic disc regions. After blurring, the coordinates of the brightest pixel are extracted, and a 512 × 512 ROI is cropped around this location. ROI extraction performance was evaluated using a localization success metric based on overlap with ground-truth optic disc regions, yielding an accuracy of 88%. Figure 4 illustrates conventional bright-spot detection with the proposed Gaussian-based approach. Diagrams show that maximum intensity decreases from 6789 before denoising to 169 after smoothing, enabling more stable ROI selection.

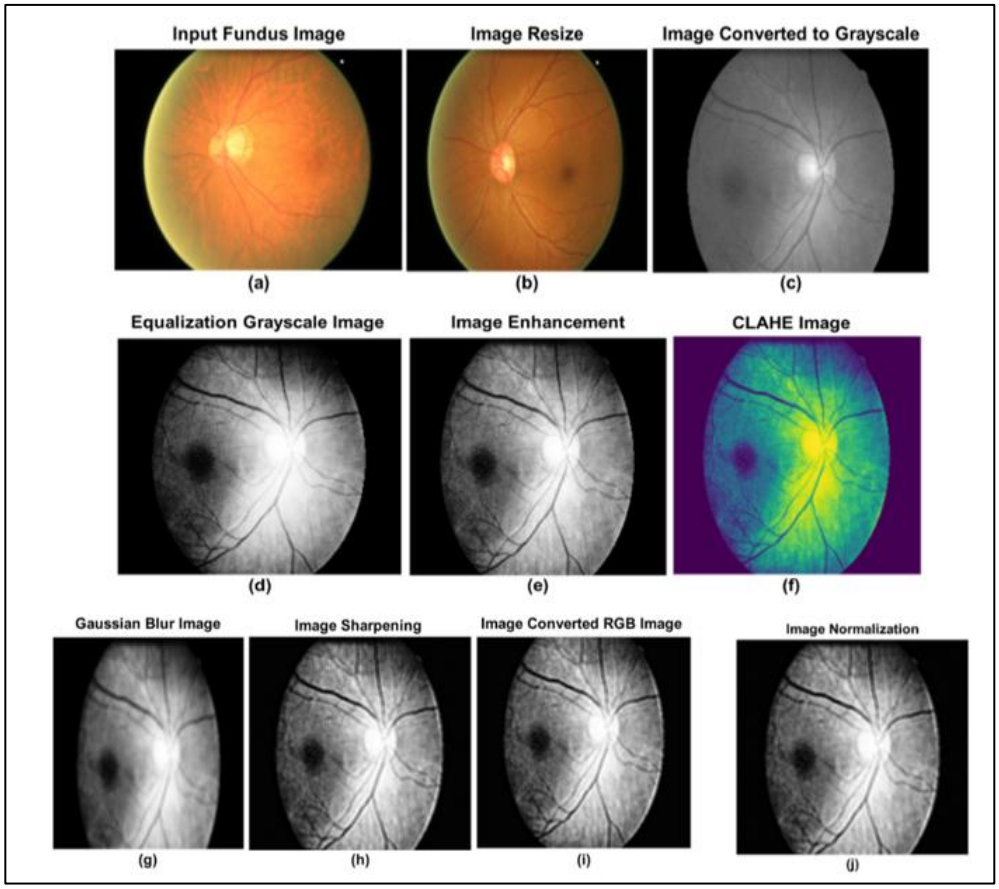

Figure 4. The Brightest Spot Technique is Used to Extract ROI. The Findings of the Conventional Method are Shown in (a – f). The Outcomes of Applying the Suggested Strategy Are Shown in (g – j)

4.3 Check Data Balancing and Apply ADASYN Technique

The oversampling technique known as ADASYN (Adaptive Synthetic Sampling) attempts to solve the problem of class imbalance by synthesizing minority class samples according to their difficulty in learning. Rather than introducing samples at an equal rate, ADASYN focuses on areas where the minority population is not dense and is closer to the majority population, particularly near the decision boundary. A difficulty value is assigned to each minority sample, with more challenging parts receiving higher numbers of synthetic points. New samples are then generated by interpolating between a minority example and a minority neighbor. This method is especially applicable to medical samples, such as those used

in glaucoma detection, where the patterns of minority classes may be subtle and therefore easily lost in the majority class. In this work, ADASYN is employed to analyze and balance the dataset by generating additional glaucoma samples, and the class distribution is visually presented before and after oversampling. ADASYN was applied exclusively in the feature space during CNN training to address class imbalance. It was not used to generate or modify medical images, optic disc or cup morphology, or segmentation masks, and no synthetic samples were included in the validation or test sets. Because ADASYN is applied only to feature representations after segmentation, it does not influence optic disc or cup morphology. A bar plot is used to show the counts and proportions of both classes non-glaucoma (label 0) and glaucoma (label 1). Such a procedure detection of detect imbalance at a very early stage, thus enabling more effective training for the classification model. ADASYN does not affect optic disc or cup morphology since it is applied after feature extraction and not on segmentation masks. ADASYN was applied only to the training dataset. Validation and test datasets were kept unchanged to avoid data leakage. In Figure 5, the original imbalanced dataset is displayed, whereas in Figure 6, the balanced dataset after ADASYN application is depicted.

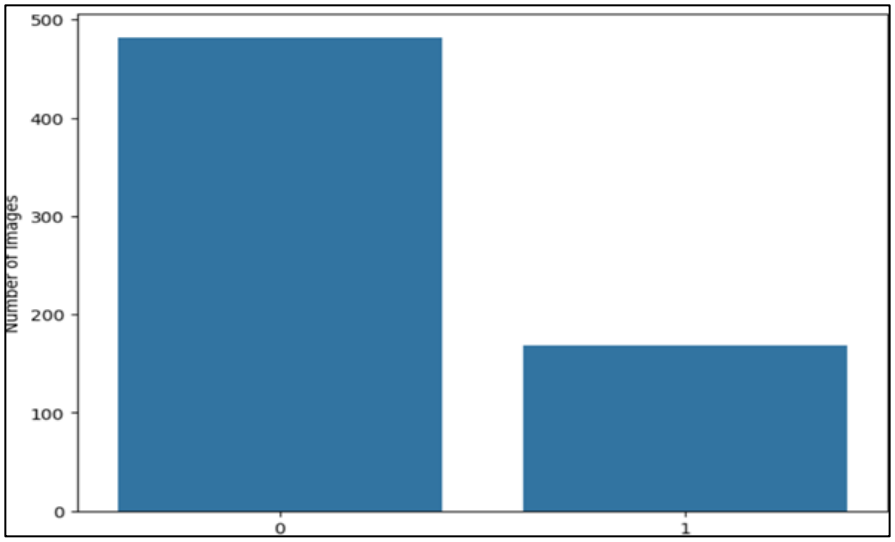

Figure 5. Data Balancing Label Counts 0 Represents 482 And 1 Represents 162 Samples with Class Proportion 0 Represents 0.741538 And 1 Represents 0.258462

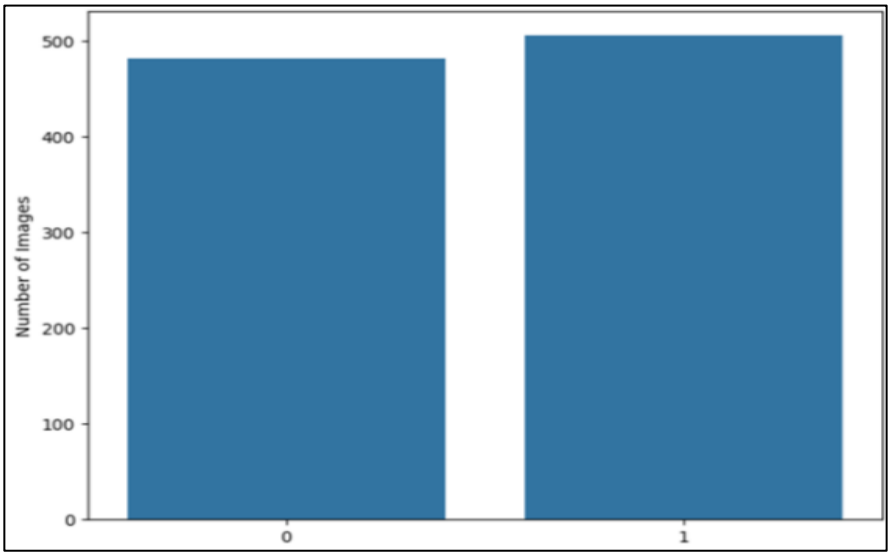

Figure 6. Applying ADASYN Technique 1 Represents 506 Samples And 0 Represents 482 Samples, Image Shape After Resembling 988 Samples

4.4 Data Pre-Processing

To maintain the same level of quality and ensure training readiness, every fundus image undergoes a standardized preprocessing pipeline. All images are scaled to 256 × 256 pixels and converted to grayscale, and Contrast Limited Adaptive Histogram Equalization (CLAHE) is applied to facilitate local contrast [21]. Histogram equalization was applied to improve global contrast and reduce illumination variability across fundus images, while CLAHE was used to enhance local contrast and highlight fine retinal structures without excessive noise amplification. The complementary use of these techniques improves the visibility of optic disc and cup regions for downstream analysis. All images were resized to 256 × 256 pixels to standardize input dimensions, reduce computational complexity, and ensure compatibility with the CNN- and U-Net–based architectures, while preserving sufficient anatomical detail. Moreover, the global contrast adjustment done through histogram equalization enhances local contrast. Gaussian filters remove high-frequency noise, and sharpening filters enhance the visibility of edges and anatomical features. At the very end, the pixel intensities are normalized to a fixed range to help CNN training stay stable. The pre-processed arrays contain the images and labels, and their shapes at the end of the pipeline are checked. Figure 7 displays a sample of the representative pre-processed samples, demonstrating that resizing, enhancement, and labelling have been performed correctly. Class names were converted into binary form before model training.

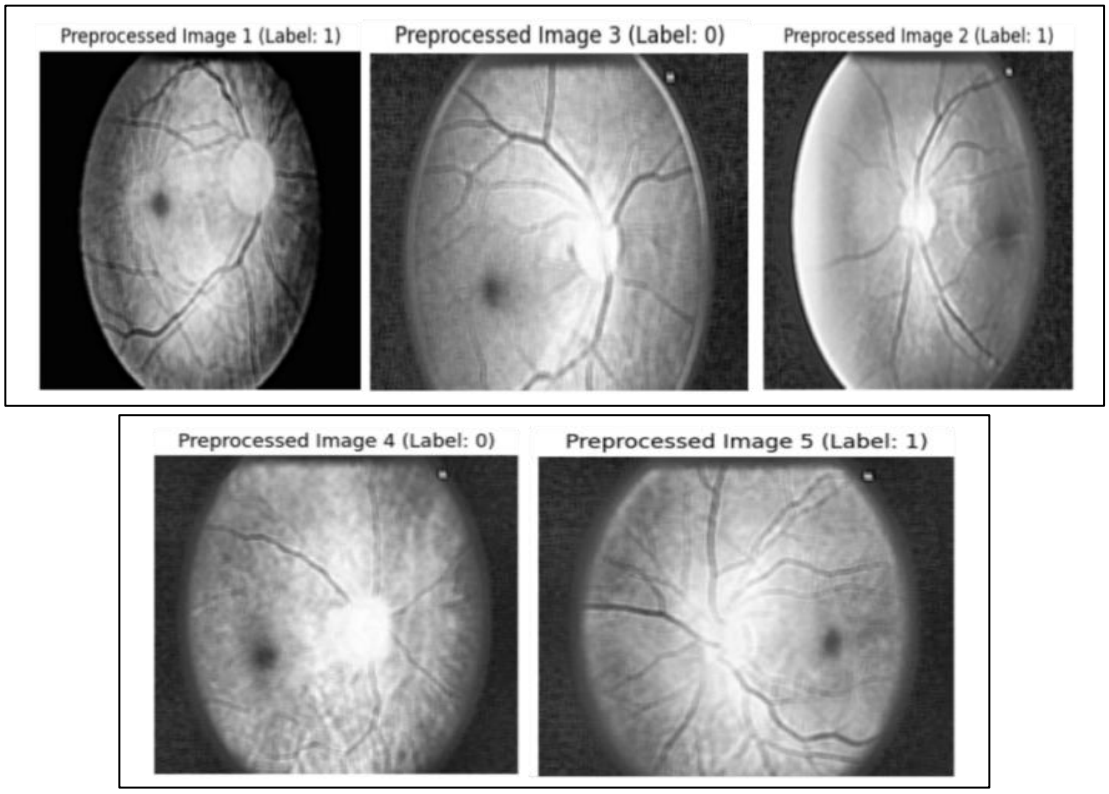

Figure 7. Visual Inspection Steps Used to Verify the Preprocessing of Training Images

4.5 Segmentation of Optic Disc and Optic Cup using U-Net

U-Net, cited in [22], is a deep convolutional neural network (CNN) that has become the most influential and popular CNN architecture in biomedical image segmentation. This is largely attributable to its symmetric encoder–decoder design, which enables it to yield

satisfactory results even with a few training samples, and its modularity, which has stimulated the rapid development of numerous architectural variants. Since its inception in 2015, the paper has been cited more than 48,000 times, a compelling indication of its effectiveness in various imaging disciplines. Essentially, the U-Net features, such as skip connections, backbone enhancements, bottleneck alterations, transformer-based extensions, and probabilistic modelling, enable it to be used for OD and OC segmentation with high accuracy.

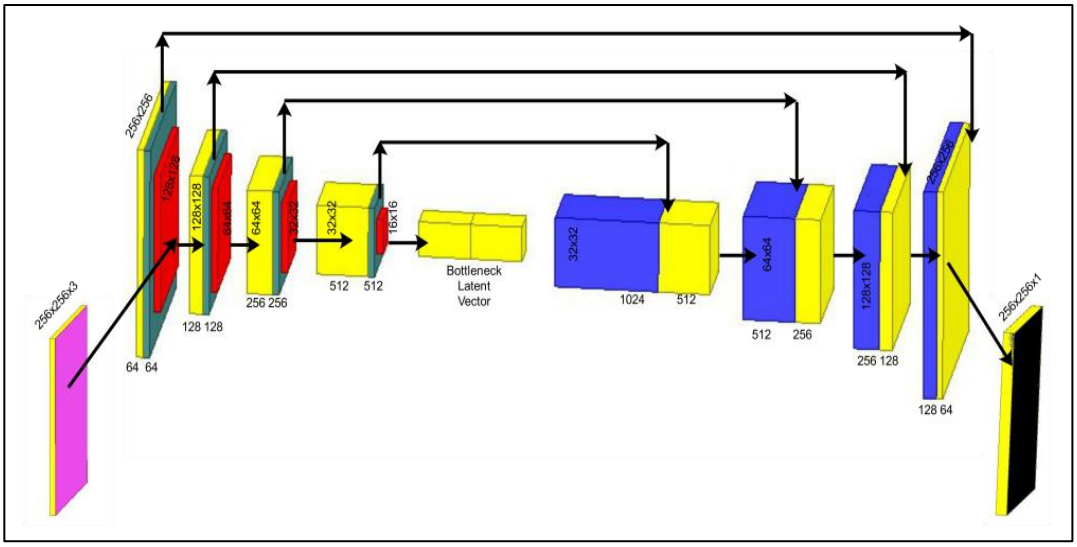

Figure 8. U-Net-Based Optic Disc and Optic Cup Segmentation

The encoder uses 3×3 convolutions, ReLU activation (ReLU(x) = max (0, x)), and 2×2 max pooling for hierarchical features. Then, the decoder upsamples and combines high- and low-level features via skip connections. Weight initialization, batch normalization, dropout (50%), and stochastic gradient descent (SGD) support the ability to train the network without crashing and its eventual generalization to unseen data. U-Net employs a non-linear convolutional encoder–decoder network and minimizes a regularized training error, with SGD updates, and it uses a Sigmoid activation to constrain output probabilities. Figure 8 provides the layer-wise functional summary of the U-Net model. After segmentation, feature extraction is performed on OD and OC regions, including computation of the clinically vital cup-to-disc ratio (CDR). Optic disc and cup segmentation is used as an intermediate step for ROI localization and cup-to-disc ratio estimation; therefore, segmentation performance is evaluated through its impact on downstream glaucoma classification rather than standalone pixel-level overlap metrics such as the Dice coefficient or Intersection over Union (IoU). CDR is obtained using segmented binary regions, where values ≥ 0.5 typically indicate glaucomatous cupping. Evaluation of performance is mostly based on typical measures that comprise accuracy, specificity, sensitivity, precision, and AUC, thereby providing a complete account of the classification system's behaviour. AUC measures the model's ability to differentiate between different classes at various decision thresholds. AUC values greater than 0.5 indicate better performance than random classification. Therefore, AUC is considered a major evaluation parameter. The standard decision boundary for balanced binary classification following ADASYN oversampling was set with an AUC value of 0.5.

4.6 Classification

The classification component enhances the robustness of the glaucoma detection system by integrating a U–Net–based OD/OC segmentation with a CNN-driven prediction. After the

U-Net extracts the optic disc and cup regions and computes the clinically significant cup-to-disc ratio (CDR), these CDR values are used as high-level diagnostic features. Essentially, a majority voting method combines CNN results with U-Net-derived CDR to produce the most stable final decision. CNNs undergo training for 150 epochs with a batch size of 36, and an 80/20 train-test split is used for proper training. This integrated pipeline—U-Net segmentation, CDR extraction, CNN classification, and majority voting—enhances both precision and clinical trustworthiness in glaucoma diagnosis.

CNNs are best suited for retinal image classification as they can automatically detect visual patterns by their nature through convolution, ReLU activation, and max pooling. The feature maps obtained are flattened and then routed to dense layers, whereas binary glaucoma prediction is done with the help of sigmoid activation. An additional MLP model (one input layer, two hidden layers, one output layer) with ReLU and sigmoid activations is trained using a categorical cross-entropy loss function. The CNN architecture used here is displayed in Figure 9. Figure 10 illustrates the complete framework of the proposal.

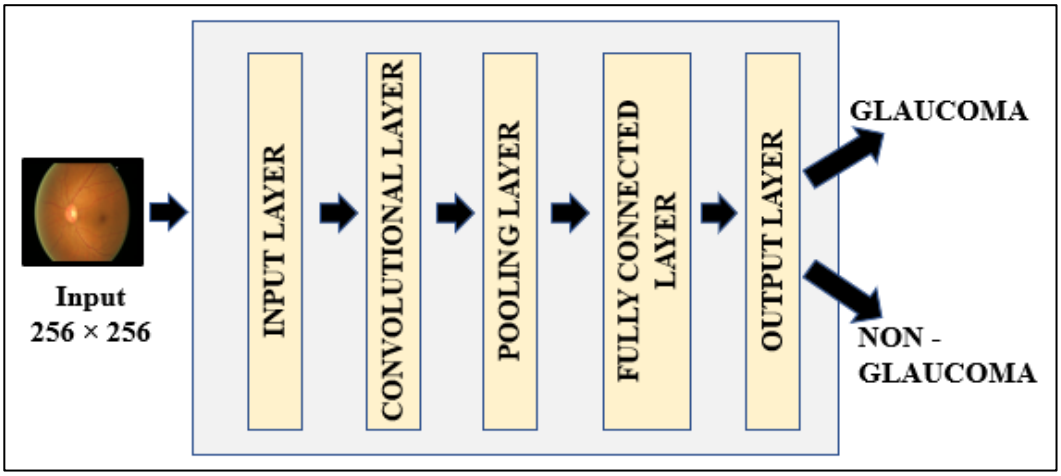

Figure 9. Core Architecture of Deep Learning

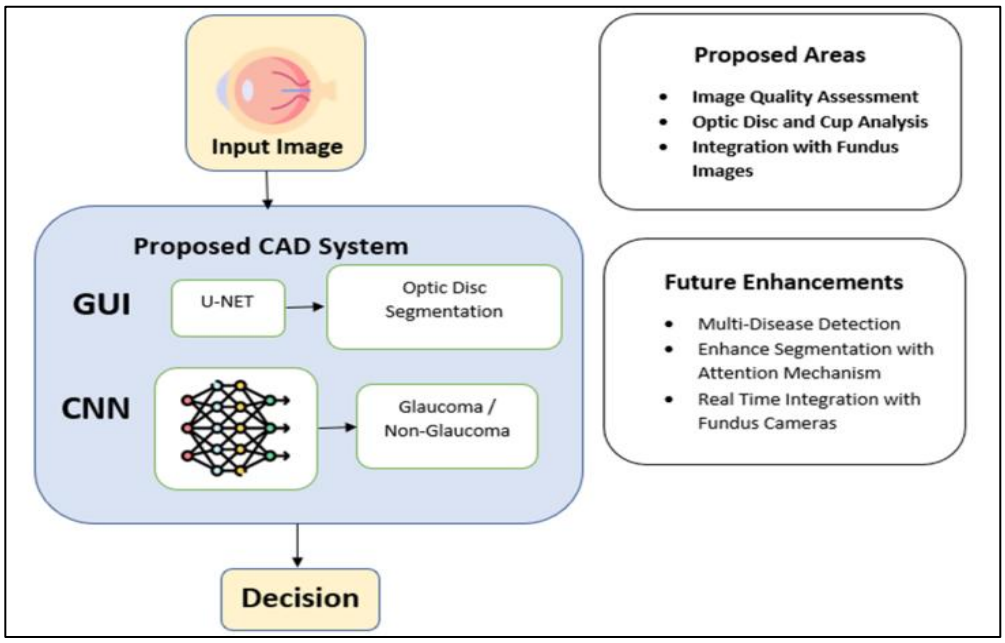

Figure 10. Overview of the Proposed CAD System for Glaucoma Detection

5. Results and Discussions

Experimental results have demonstrated U-Net to be a very effective method for OD/OC segmentation, as it obtained the following metric values: accuracy 86.38%, sensitivity 88.95%, specificity 85.43%, precision 84.77%, and AUC 91.63%. Training was performed on 256 × 256 resized fundus images using stochastic gradient descent with a momentum of 0.95 and a fixed learning rate of 0.0004. The model was trained for 150 epochs with a batch size of 16, employing an 80/20 train–test split. Implementation was done in Python 3.7 using TensorFlow and Keras on an NVIDIA Tesla GPU with CUDA 10.1. Figures 11 14 illustrate the segmentation outputs, confusion matrix, performance curves, and AUC. For optic disc and optic cup segmentation, pixel-level AUC was computed using the probability maps generated by the U-Net model and the corresponding ground-truth segmentation masks. In this case, AUC measures the model's ability to distinguish foreground (disc/cup) pixels from background pixels across all thresholds. The segmentation-based AUC is 0.9163.

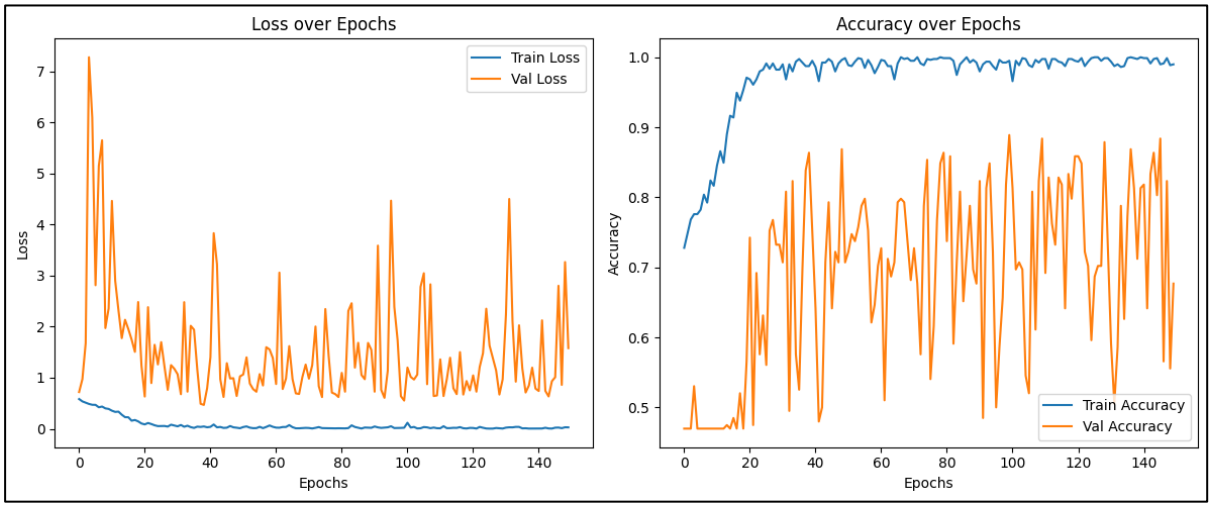

Figure 11. Performance Graph in Loss Over Epochs and Accuracy Over Epochs

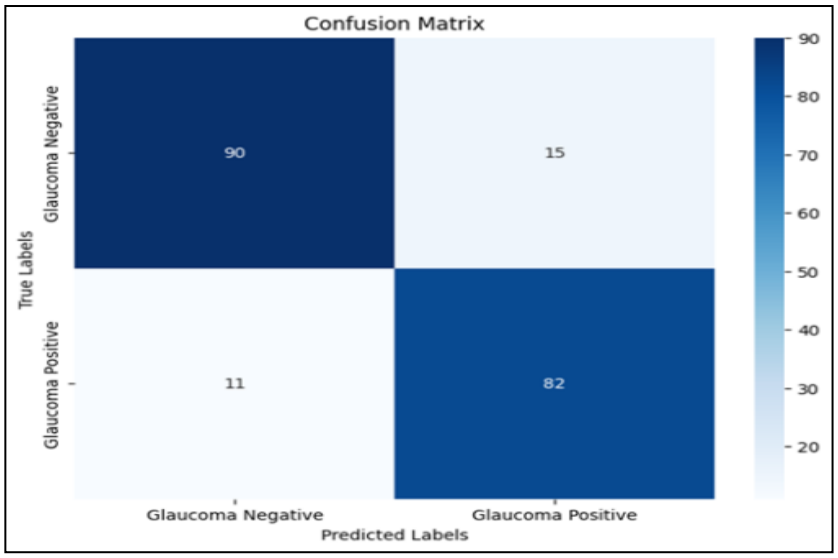

Figure 12. Confusion Matrix of U-Net Model

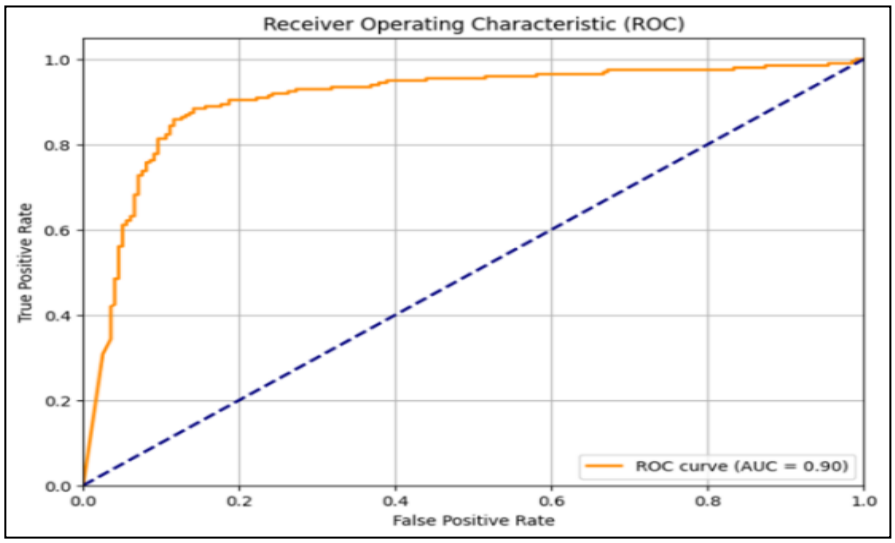

Figure 13. Comparison of Area Under Curve (AUC)

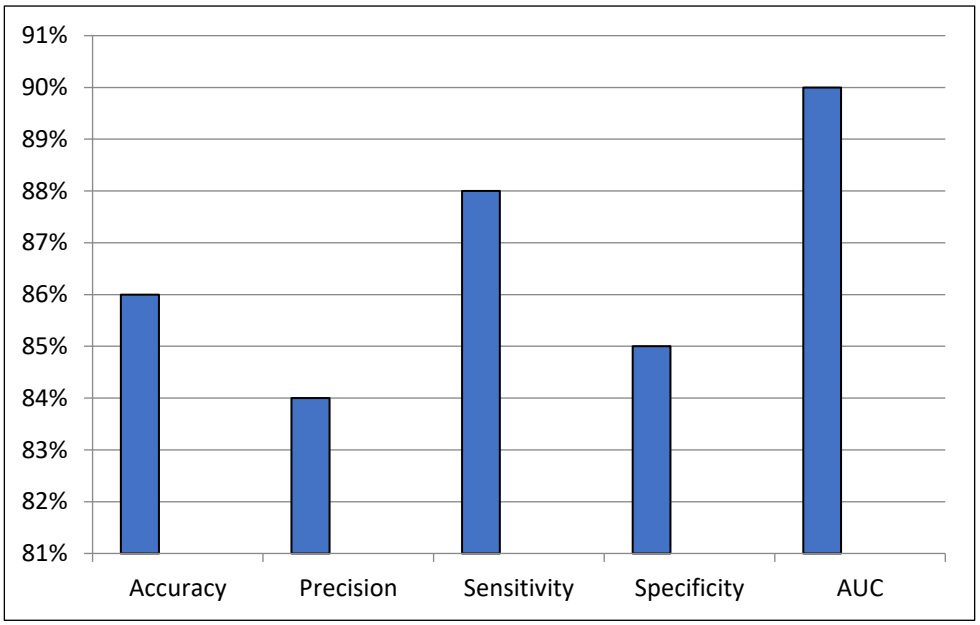

Figure 14. Classification Networks: Accuracy, Precision, Sensitivity, Specificity, AUC

In this section, the results of the proposed CAD pipeline are presented for different main stages, including input image validation, segmentation, feature extraction, and classification. The LeNet model was applied for checking the validity of input images, and it showed stable training and validation loss curves, which indicates that the model was trained properly and no overfitting occurred. Accurate segmentation of the optic disc (OD) and optic cup (OC) is important because the cup-to-disc ratio (CDR) is the main indicator used for glaucoma detection. However, conventional image processing methods face difficulties due to changes in illumination, image resolution, and noise. To address these issues, the U-Net model was used for optic disc and cup segmentation. With preprocessing enhancements such as CLAHE, Gaussian smoothing, dropout, and data augmentation, the U-Net produced reliable OD/OC segmentations with significantly reduced prediction time. The performance metrics and comparative analysis are summarized in Table 1 in terms of Accuracy, Sensitivity, Specificity, Precision, and AUC. The results also suggest that bigger training datasets lead to higher classification accuracy. to illustrate the model's performance.

Table 1. U-Net Outcome

Classifier	Accuracy	Specificity	Sensitivity	Precision	AUC
U-Net	86.38%	85.43%	88.95%	84.77%	91.63%

After integrating U-Net and CNN into the classification process, accuracy and precision percentage have improved substantially. As shown in Figure 15, the confusion matrix contains all True Positives (TP), True Negatives (TN), False Positive (FP) and False Negative (FN). All these results indicate that the proposed approach is very reliable and efficient for glaucoma detection and diagnosis of glaucoma by deep learning. Table 2 demonstrates the success of applying CNN on U-Net output. The accuracy values that are less than those in Table 2 suggest a human set that is conservative or unbiased. Through the training of the models, they did not utilize transfer learning or extensive hyperparameter tuning nor did they add any aggressive data augmentations. The dataset was limited. Thus, they provide a realistic view of the performance of generalization. The loss and accuracy curves against epochs and AUC metrics are shown from Figure 16 to Figure 18. The predicted probability scores produced by the CNN classifier and their corresponding ground-truth labels on the held-out test set are used to compute AUC for glaucoma classification. The ROC curve plots the true positive rate versus the false positive rate for various threshold values. The AUC indicates the classification score difference between a random glaucoma image and a random non-glaucoma image, and thus quantifies the classifier's discriminative power independent of any threshold selected. CNN-based AUC is 0.9324.

Table 2. U-Net + CNN Classifier Outcome

Classifier	Accuracy	Specificity	Sensitivity	Precision	AUC
U-Net + CNN	88.38%	81.43%	84.95%	89.77%	93.24%

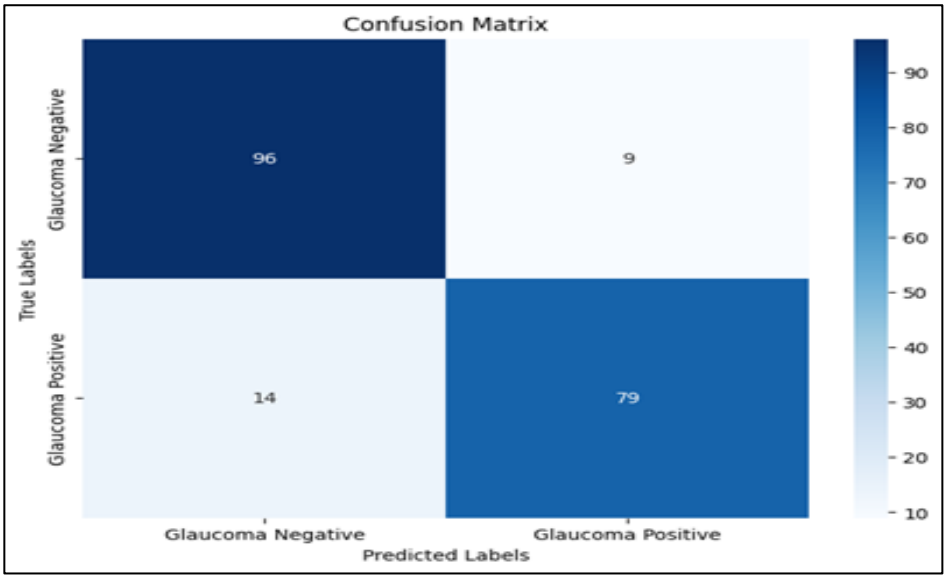

Figure 15. Confusion Matrices

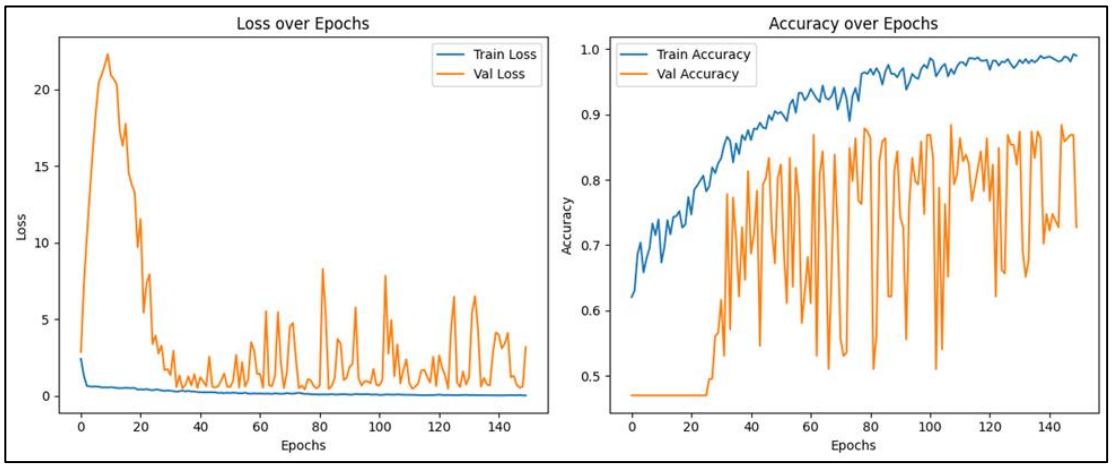

Figure 16. Performance Graph in Loss Over Epochs and Accuracy Over Epochs

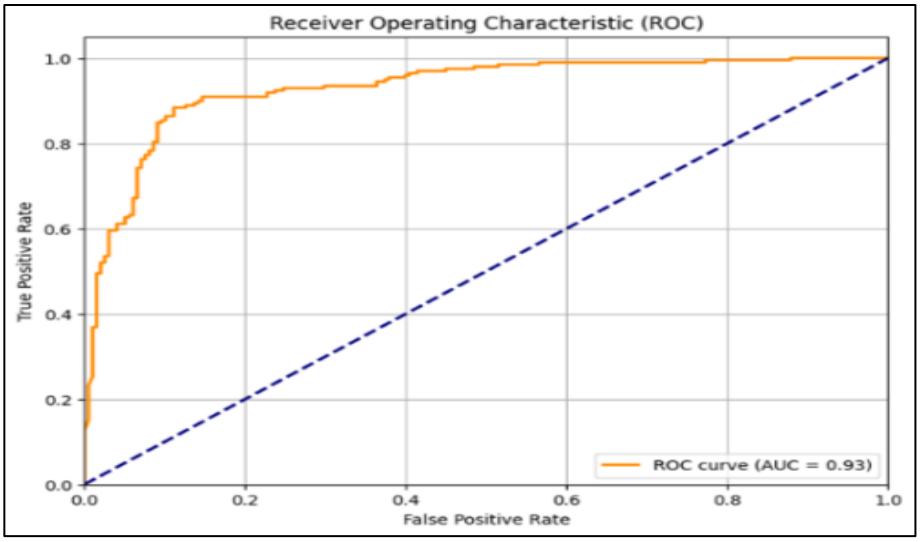

Figure 17. Comparison of Area under Curve (AUC)

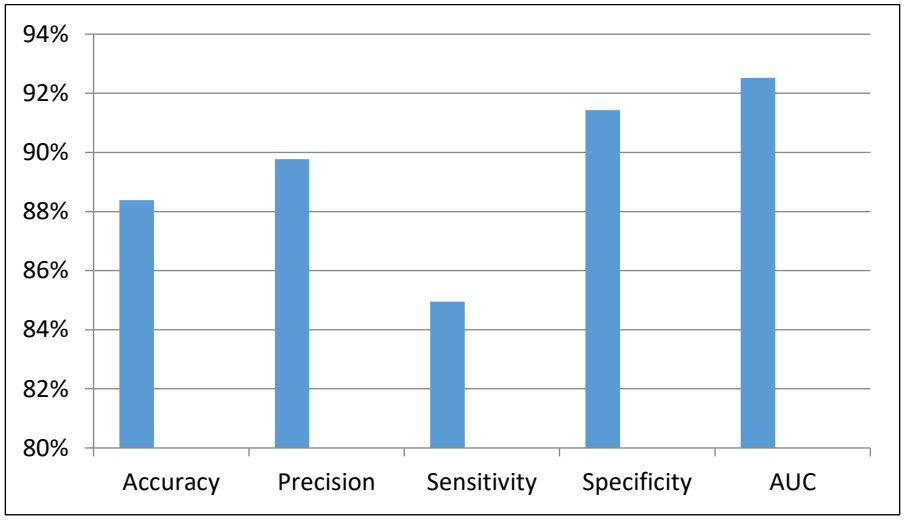

Figure 18. Classification Networks: Accuracy, Precision, Sensitivity, Specificity, AUC

Many of the misclassified cases have some common failure modes. Most mistakes mainly occur in patients with early glaucoma. The optic nerve head modifications due to glaucoma are small. Moreover, the cup-to-disc ratio is close to the decision threshold. The outcomes of both optical disc and cup segmentation and classification are affected by poor

image quality like low contrast and non-uniform illumination. Normal anatomical variations can produce false positive results such as larger optic cups in non-glaucoma eyes. In addition, minor errors in the segmentation of disc and cup can slip through the process.

5.1 Comparison with Existing Work

Table 3 compares state-of-the-art glaucoma detection frameworks. A direct comparison between them is not entirely justified as they are trained and evaluated on different datasets with different characteristics, imaging quality, class distribution and annotation protocols. Nonetheless, the comparison provides a useful context for illustrating the pattern of existing methods. Amidst this wide array of results, the proposed scheme exhibits a well-balanced performance with an accuracy of 88.38%, specificity of 81.43%, sensitivity of 84.95%, precision of 89.77% and an AUC of 93.24%. The Transfer Learning + U-Net method proposed in [7] achieves a higher accuracy rate than that of the current model. However, a raw comparison of metrics is often inaccurate as datasets are inherently variable. Nonetheless, the proposed model shows a robust balance on most clinical indices, hinting at its practical use for glaucoma prediction.

Table 3. Performance Comparison of Deep Learning Algorithms for Glaucoma Detection

Ref.	Classifier	Accuracy	Specificity	Sensitivity	Precession	AUC
[23]	ConvNeXt_Base + CBAM + Meta	81.77%	85.92%	76.60%	-	89.3%
[24]	KNN + CNN	-	74%±09%	81%±15%	-	78%±11%
[7]	Transfer Learning + U-Net	96.33%	98.15%	97.03%	-	-
[25]	Logistic Regression + Deep Learning	96%	93%	100%	-	98.0%
This work	U-Net + CNN	88.38%	81.43%	84.95%	89.77%	93.24%

5.2 Display a Test Image with Prediction Result

This part of the system visualizes how the binary classification model performs on the test images. The trained model first predicts class probabilities for each image in x_test, producing a value between 0 and 1. We apply a fixed threshold of 0.5, which means probabilities greater than 0.5 are classified as class 1 (i.e., positive for glaucoma), while probabilities less than or equal to 0.5 are labeled as class 0 (i.e., negative). The predictions are compiled into a single array called y-pred-labels. At the beginning of performance testing, a random selection of test images (generally around 10) is taken and displayed in a grid of 5 columns by 2 rows using Matplotlib. The title of each subplot displays both the actual and the predicted label. The color of the title indicates the prediction status: green indicates a correct prediction; red indicates an incorrect assignment. These visualizations help identify which samples have been classified correctly and which have not. They act as debugging tools for performance and also help visualize the results. Figure 19 displays the test images and CNN predictions, which are similar to those used during training.

5.3 Positive (Glaucoma) and Negative (Healthy) Prediction Using the Streamlit App

The final stage of the Streamlit app developed glaucoma detection system shows users the prediction result along with the associated retinal fundus image for easy visual understanding. The uploaded samples will be referred to for the construction of one negative image and one positive image. The model labelled image 041.jpg (image number) as Positive

(Glaucoma) with a probability/confidence score of 0.77. This means that the abnormal features of the optic disc and the neighbouring retina being used for the diagnosis are very likely those consistent with glaucomatous damage. The disease can be predicted with the help of clinical indicators like the bigger cup-to-disc ratio, loss of the nerve fiber layer, and thinning of the neuroretinal rim. This type of visual and quantitative feedback can be very useful for screening and triaging patients. The image, 598.jpg (image no), at the same time, has a very high confidence score of 0.99, showing very great model confidence the eye is healthy and is predicted to be Negative (No Glaucoma). The diagram below illustrates a normal retina, characterized by an optic disc having symmetrical branching of the vasculature, a small cup and distinct disc borders. To ensure that the application remains responsive and accurately scaled, st.image() is used to present the predictions in the application, alongside the option use_container_width = True. This interface is a standard tool for researchers, students, and ophthalmologists to visually and numerically confirm categorizations, which in turn helps build trust in the clinical examination of AI-generated results. Figure 20 shows the Streamlit framework for presenting the CAD tool for glaucoma detection.

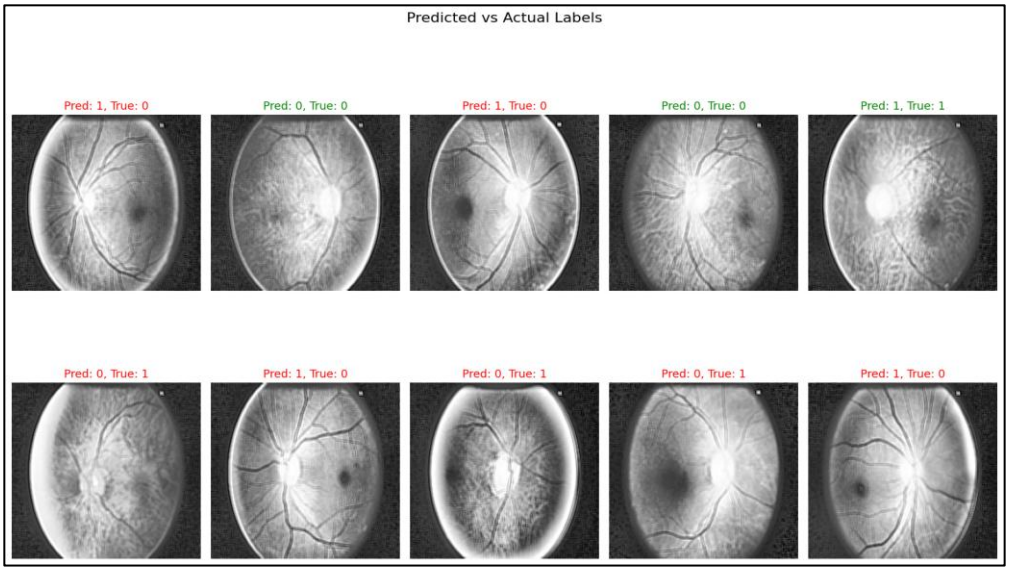

Figure 19. Displaying the Test Image by Applying CNN with the Prediction Result

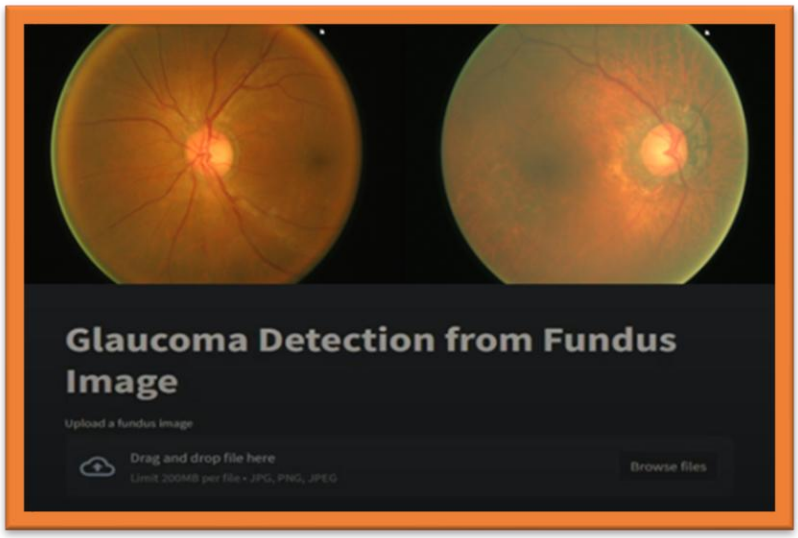

Figure 20. Streamlit Framework for Presenting the CAD Tool for Glaucoma Detection

6. Conclusion

The proposed model's novel hybrid architecture consists of a U-Net that efficiently segments the optic disc (OD) and optic cup (OC), and a CNN classifier that diagnoses glaucoma. Combining the two methods allows for obtaining a clinically significant cup-to-disc ratio (CDR) with a high level of certainty, while simultaneously detecting the ailment using both structural and learnt properties seamlessly. The U-Net system produced sufficient segmentation results with an accuracy of 86.38%, sensitivity of 88.95%, specificity of 85.43%, precision of 84.77%, and AUC of 91.63%. The diagnostic reliability was further improved with the CNN classifier, which achieved the highest accuracy, sensitivity, specificity, precision, and AUC of 88.38%, 84.95%, 81.43%, 89.77%, and 93.24%, respectively. The results suggest that the proposed method can differentiate between healthy and glaucomatous eyes across different data sets. One of the major advantages of the system is its effective usability. A doctor with limited technological expertise can upload images, visually check the output segmentation, and obtain a prediction through the graphic interface built with Streamlit without much difficulty. This tool can be used in all clinical departments, as well as in remote tele-ophthalmology units, which can apply it easily and effectively. It is possible to implement this procedure as an early training step. Further work should aim to enhance detection and confirm its efficacy in other clinics. Lastly, the CAD system is an effective, accurate, and useful AI tool that can be widely utilized in the initial stages of glaucoma screening.

Conflicts of Interest/Competing Interests

The authors declare that there is no conflicts of interest/competing of interests.

Funding Information

No fund received for this work

Ethical Statement

The present study does not involve human subjects or animal models. Hence, no ethical approval was necessary.

Acknowledgements

The authors have no acknowledgements to declare.

References

[1] Yadav, Vaibhav, Barnali Dey, Udayan Baruah, Saumya Das, and Om Prakash. "Machine Learning-Assisted Image Analysis Techniques for Glaucoma Detection." EURASIP Journal on Image and Video Processing 2025, no. 1 (2025): 7.

[2] Thakur, Niharika, and Mamta Juneja. "Early-Stage Prediction of Glaucoma Disease to Reduce Surgical Requirements Using Deep-Learning." Materials today: proceedings 45 (2021): 5660-5664.

[3] Ajitha, S., John D. Akkara, and M. V. Judy. "Identification of Glaucoma from Fundus Images Using Deep Learning Techniques." Indian journal of ophthalmology 69, no. 10 (2021): 2702-2709.

[4] Li, Feng, Lei Yan, Yuguang Wang, Jianxun Shi, Hua Chen, Xuedian Zhang, Minshan Jiang, Zhizheng Wu, and Kaiqian Zhou. "Deep Learning-Based Automated Detection of Glaucomatous Optic Neuropathy on Color Fundus Photographs." Graefe's Archive for Clinical and Experimental Ophthalmology 258, no. 4 (2020): 851-867.

[5] Hussain, Rukhshanda, and Hritam Basak. "Ut-Net: Combining U-Net and Transformer for Joint Optic Disc and Cup Segmentation and Glaucoma Detection." arXiv preprint arXiv:2303.04939 (2023).

[6] Ran, An Ran, Clement C. Tham, Poemen P. Chan, Ching-Yu Cheng, Yih-Chung Tham, Tyler Hyungtaek Rim, and Carol Y. Cheung. "Deep Learning in Glaucoma with Optical Coherence Tomography: A Review." Eye 35, no. 1 (2021): 188-201.

[7] Kashyap, Ramgopal, Rajit Nair, Syam Machinathu Parambil Gangadharan, Miguel Botto-Tobar, Saadia Farooq, and Ali Rizwan. "Glaucoma Detection and Classification Using Improved U-Net Deep Learning Model." In Healthcare, vol. 10, no. 12, MDPI, 2022, 2497.

[8] Helman, Stephanie, Martha Ann Terry, Tiffany Pellathy, Andrew Williams, Artur Dubrawski, Gilles Clermont, Michael R. Pinsky, Salah Al-Zaiti, and Marilyn Hravnak. "Engaging Clinicians Early During the Development of a Graphical User Display of an Intelligent Alerting System at the Bedside." International journal of medical informatics 159 (2022): 104643.

[9] Ali, Aziah, Wan Mimi Diyana Wan Zaki, Aini Hussain, Noramiza Hashim, and Wan Noorshahida Mohd Isa. "Vessel Masking and Hough Transform for Optic Disc Localisation from Retinal Images." F1000Research 11 (2022): 181.

[10] Toptaş, Buket, Murat Toptaş, and Davut Hanbay. "Detection of Optic Disc Localization from Retinal Fundus Image Using Optimized Color Space." Journal of Digital Imaging 35, no. 2 (2022): 302-319.

[11] Veena, H. N., A. Muruganandham, and T. Senthil Kumaran. "A Novel Optic Disc and Optic Cup Segmentation Technique to Diagnose Glaucoma Using Deep Learning Convolutional Neural Network Over Retinal Fundus Images." Journal of King Saud University-Computer and Information Sciences 34, no. 8 (2022): 6187-6198.

[12] Dhar, Sanchari, and Lior Shamir. "Evaluation of the Benchmark Datasets for Testing the Efficacy of Deep Convolutional Neural Networks." Visual informatics 5, no. 3 (2021): 92-101.

[13] Aamir, Muhammad, Muhammad Irfan, Tariq Ali, Ghulam Ali, Ahmad Shaf, Alqahtani Saeed S, Ali Al-Beshri, Tariq Alasbali, and Mater H. Mahnashi. "An Adoptive Threshold-Based Multi-Level Deep Convolutional Neural Network for Glaucoma Eye Disease Detection and Classification." Diagnostics 10, no. 8 (2020): 602.

[14] Munira, Hafiza Akter, and Md Saiful Islam. "Hybrid Deep Learning Models for Multi-Classification of Tumour from Brain MRI." Journal of Information Systems Engineering and Business Intelligence 8, no. 2 (2022): 162-174.

[15] Yang, Cheng, Sung-Kwun Oh, Bo Yang, Witold Pedrycz, and Lin Wang. "Hybrid Fuzzy Multiple SVM Classifier Through Feature Fusion Based on Convolution Neural Networks and Its Practical Applications." Expert Systems with Applications 202 (2022): 117392.

[16] Alharbi, Meshal. "Multi-Classification of Eye Disease Based on Fundus Images Using Hybrid Squeeze Net and LRCN Model." Multimedia Tools and Applications 83, no. 27 (2024): 69197-69226.

[17] Das, Dipankar, Deepak Ranjan Nayak, Sulatha V. Bhandary, and U. Rajendra Acharya. "CDAM-Net: Channel shuffle dual attention based multi-scale CNN for efficient glaucoma detection using fundus images." Engineering Applications of Artificial Intelligence 133 (2024): 108454.

[18] Zakariah, Mohammed, Salman A. AlQahtani, and Mabrook S. Al-Rakhami. "Machine Learning-Based Adaptive Synthetic Sampling Technique for Intrusion Detection." Applied Sciences 13, no. 11 (2023): 6504.

[19] Khan, Umer Sadiq, and Saif Ur Rehman Khan. "Boost Diagnostic Performance in Retinal Disease Classification Utilizing Deep Ensemble Classifiers Based on OCT." Multimedia Tools and Applications 84, no. 19 (2025): 21227-21247.

[20] Guru Prasad, M. S., H. N. Naveen Kumar, K. Raju, D. K. Santhosh Kumar, and S. Chandrappa. "Glaucoma Detection Using Clustering and Segmentation of the Optic Disc Region from Retinal Fundus Images." SN computer science 4, no. 2 (2023): 192.

[21] Mohammed, Ibrahim Majid, and Nor Ashidi Mat Isa. "Contrast Limited Adaptive Local Histogram Equalization Method for Poor Contrast Image Enhancement." IEEE Access (2025).

[22] Azad, Reza, Ehsan Khodapanah Aghdam, Amelie Rauland, Yiwei Jia, Atlas Haddadi Avval, Afshin Bozorgpour, Sanaz Karimijafarbigloo, Joseph Paul Cohen, Ehsan Adeli, and Dorit Merhof. "Medical Image Segmentation Review: The success of U-Net." IEEE Transactions on Pattern Analysis and Machine Intelligence (2024).

[23] Chiang, Yen-Ying, Ching-Long Chen, and Yi-Hao Chen. "Deep Learning Evaluation of Glaucoma Detection Using Fundus Photographs in Highly Myopic Populations." Biomedicines 12, no. 7 (2024): 1394.

[24] Sánchez-Morales, Adrián, Juan Morales-Sánchez, Oleksandr Kovalyk, Rafael Verdú-Monedero, and José-Luis Sancho-Gómez. "Improving Glaucoma Diagnosis Assembling Deep Networks and Voting Schemes." Diagnostics 12, no. 6 (2022): 1382.

[25] Akter, Nahida, John Fletcher, Stuart Perry, Matthew P. Simunovic, Nancy Briggs, and Maitreyee Roy. "Glaucoma Diagnosis Using Multi-Feature Analysis and a Deep Learning Technique." Scientific Reports 12, no. 1 (2022): 8064.

[26] https://www.kaggle.com/datasets/sshikamaru/glaucoma-detection