Automated Glaucoma Detection from Retinal Fundus Images Using Deep Learning

Aarya Nanda Sajan

A report submitted as part of the requirements for the Internship
Computer Science and Engineering
NIT Calicut

June 2025

Supervisor Dr. Jayaraj P B

Abstract

Glaucoma is a progressive optic neuropathy and a major cause of irreversible blindness. Accurate detection at an early stage is essential for preventing vision loss. While traditional methods rely heavily on cup-to-disc ratio (CDR), this measure often fails to capture finer structural changes in the optic nerve head. In this work, we present a hybrid deep learning-based model that incorporates both anatomical and geometric features for glaucoma detection. Fundus images are first filtered using FundusQ-Net to ensure quality, followed by preprocessing and vessel segmentation. The model then estimates neuroretinal rim thickness and optic boundaries using both segmentation masks and the bending patterns of retinal blood vessels—an anatomical indicator used by clinicians to identify cup margins. This combination allows for a more precise and structure-aware identification of glaucomatous damage. By integrating vessel curvature analysis with neuroretinal rim evaluation, the proposed method aims to improve the interpretability, accuracy, and clinical relevance of automated glaucoma screening systems.

Acknowledgements

I would like to express my deepest gratitude to everyone who supported me throughout this internship.

Firstly, I am profoundly thankful to my supervisor, Dr. Jayaraj P B, for their invaluable mentorship, guidance, and constant encouragement. Their expertise and constructive feedback were instrumental in shaping my learning experience and ensuring the success of my project. I sincerely thank Dinoy, my PhD scholar mentor, for his guidance in research, support with literature reviews, and valuable insights into deep learning and medical image analysis. I also extend my gratitude to the entire team at Department of Computer Science and Engineering for fostering a collaborative and supportive environment.

I'm grateful to NIT Calicut for granting me this opportunity to apply my academic knowledge in a real-world setting, and for the foundational skills I acquired prior to this internship .

Chapter 1

Introduction

Glaucoma is a group of chronic, progressive optic neuropathies characterized by the degeneration of retinal ganglion cells (RGCs) and their axons, leading to structural damage to the optic nerve head and corresponding visual field loss. The primary pathological feature of glaucoma is the gradual loss of the neuroretinal rim tissue at the optic disc, often accompanied by an increase in the cup-to-disc ratio (CDR). One of the major risk factors for glaucoma is elevated intraocular pressure (IOP), which results from impaired drainage of aqueous humor through the trabecular meshwork and Schlemm's canal. However, glaucoma can also occur in individuals with normal IOP levels, a condition known as normal-tension glaucoma. Risk factors include advanced age, family history, high myopia, diabetes, and certain ethnicities such as African, Asian, and Hispanic populations.

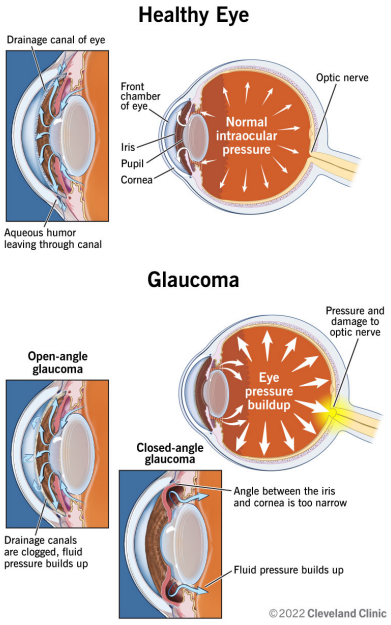


Figure 1.1: Normal eyes and Glaucomic eyes

In its early stages, glaucoma may not cause any symptoms. Symptoms may not appear until this condition causes irreversible damage. Some of the more common glaucoma symptoms include eye pain or pressure, headaches, red or bloodshot eyes, double vision (diplopia), blurred vision, gradually developing low vision, gradually developing blind spots (scotomas) or visual field defects like tunnel vision

These symptoms reflect the underlying structural damage to the optic nerve, particularly the thinning of the neuroretinal rim and the expansion of the optic cup. Traditional clinical methods rely on measurements like cup-to-disc ratio (CDR), visual field testing, and intraocular pressure readings. However, these approaches are often subjective, time-consuming, and less sensitive to early structural changes.

With the advent of retinal imaging and artificial intelligence, automated glaucoma screening systems now offer the potential for faster, more accurate, and scalable detection. This work explores current deep learning-based approaches and proposes a hybrid model that integrates anatomical and geometric features—including neuroretinal rim thickness and retinal vessel bending patterns—to enhance the accuracy and interpretability of glaucoma detection.

1.1 Conclusion and Motivation for Next Steps

In this chapter, we discussed the clinical background of glaucoma, including its causes, symptoms, risk factors, and diagnostic biomarkers. Traditional glaucoma diagnosis methods, such as measuring intraocular pressure, evaluating the cup-to-disc ratio, and conducting visual field tests, often rely heavily on the expertise and subjective judgment of the ophthalmologist. These methods can be time-consuming, prone to inter-observer variability, and may fail to detect early-stage glaucoma accurately. The mood, experience level, and fatigue of the clinician can sometimes influence the diagnostic outcome, leading to inconsistencies in diagnosis.

These limitations highlight the need for automated, objective, and more accurate diagnostic systems. The emergence of artificial intelligence and deep learning-based approaches offers promising solutions to overcome these challenges, enabling faster, scalable, and more reliable glaucoma detection. The following chapters will explore these advanced techniques in greater detail.

Chapter 2

Background

This chapter provides some background research on the project and examines some previous work.

2.1 Automated Blood Vessel Kink Detection

Wong et al. 2009

The proposed architecture for automated blood vessel kink detection focuses on using kinks as a physiological marker for accurate optic cup boundary localization in retinal images. The system starts with optic disc segmentation using a variational level set method, followed by preliminary cup boundary estimation based on pallor intensity. Localized 128×128 pixel patches are then extracted around the disc area to better capture small vessel features while minimizing the dominance of larger vessels. Vessel edges are detected using a combination of Canny edge detection and Gabor wavelet transform, both applied on the green channel of the retinal image. These edge and wavelet features are statistically fused using a Bayesian probability model to produce a more reliable vessel edge map. The edges are then vectorized and smoothed using polynomial fitting to minimize noise artifacts. Kinks are automatically detected at points where the angular change along vessel edges exceeds 20°, with additional filtering to avoid confusing curved vessels with actual kinks. An ellipse fitting process is used to generate the final cup boundary based on the detected kinks, and when kink distribution is sparse, pallor-based points are also incorporated. Performance evaluation was conducted on 27 high-resolution retinal images from the Singapore Eye Research Institute, with expert-clinician-graded cup-to-disc ratios serving as the ground truth. The kink-based method (KBM) achieved a lower RMS error (0.093) compared to the traditional pallor-based method (PBM) with 0.139, representing a 33% reduction in

error. The KBM showed significantly higher sensitivity (0.813) than PBM (0.186) for identifying high-risk glaucoma cases with CDR greater than 0.6, though it exhibited lower specificity (0.455 vs 0.818). ROC curve analysis further confirmed that KBM outperformed PBM in risk classification. The method maintained error levels within 0.2 CDR units, aligning with clinical intra-observer variability limits. KBM also reduced errors significantly at higher CDR ranges, making it valuable for detecting advanced glaucomatous changes. One limitation was the uneven spatial distribution of kinks in some images, which required supplementing points from the pallor-based boundary. Overall, the architecture's reliance on vessel morphology instead of just intensity-based features provides physiological validation of the cup boundary and makes the method robust for non-stereographic retinal images. The system shows promise as an effective tool for early glaucoma screening and optic cup segmentation.

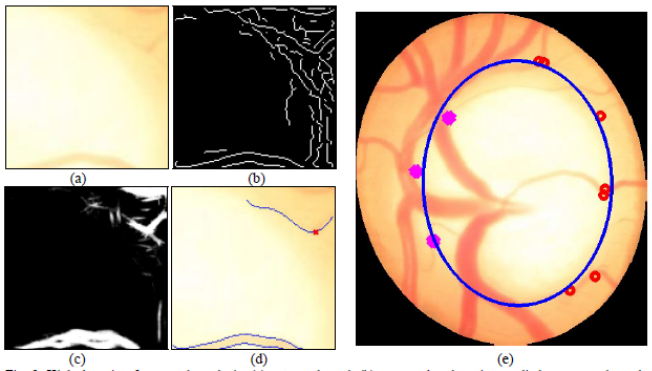


Fig. 3. Kink detection from patch analysis; (a) extracted patch (b) canny edge detection applied on green channel of the patch; (c) wavelet transform on green channel of the patch; (d) original image with superimposed detected blood vessels (blue lines) and kink (red cross) and (e) detected cup boundary based on kinks (red points) and additional points (magenta) obtained from the p pallor-based contour.

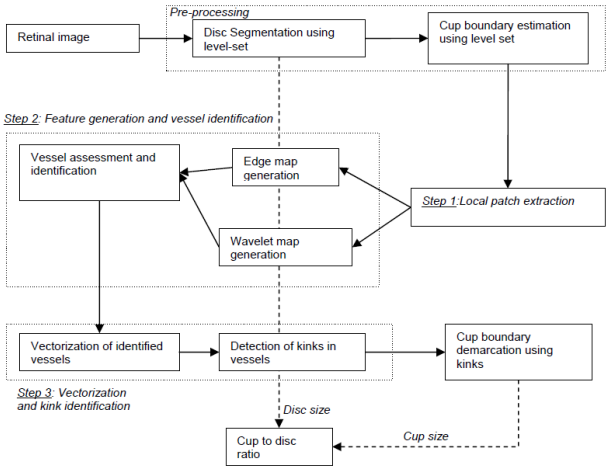


Fig. 1. Framework for the proposed system for automated kink detection to determine cup to disc ratio for assessment of glaucoma

2.2 Vessel Bend-Based Cup Segmentation in Retinal Images

Joshi, Sivaswamy, and Krishnadas 2010

The proposed method focuses on detecting the optic cup boundary from monocular color fundus images using anatomical cues like vessel bends. The approach starts with optic disc (OD) segmentation, which is done by active contour models on the red channel of the fundus image. After OD localization, the method extracts blood vessels by modeling them as trenches in a 3D intensity surface, making the process more robust to inter-image variations. The next step involves detecting vessel bends, where each bend point is called a candidate "bending point" or "bi". A Region of Support (ROS) is dynamically selected for each bend point to adaptively analyze both thick and thin vessels. The curvature profile of each vessel segment is analyzed to find local maxima, which correspond to high-curvature bending points.

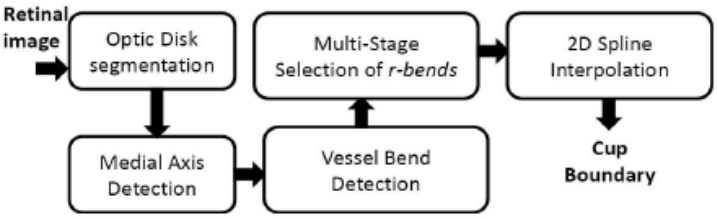


Fig. 2: The proposed method

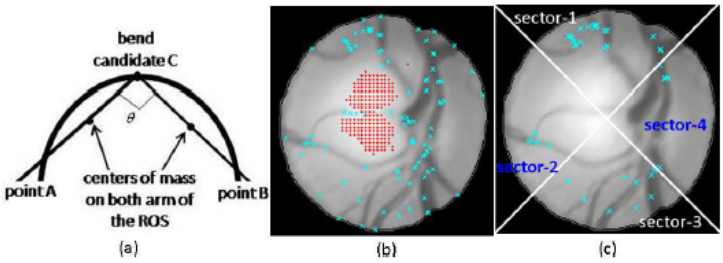


Fig. 3: a) Angle of a vessel bend, b) uniform pallor samples(red), bend points(cyan) and c) fitted circle(red) and potential r-bends.

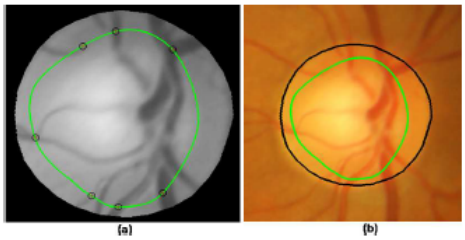


Fig. 4: a) Estimated cup boundary, b) final OD and cup boundary.

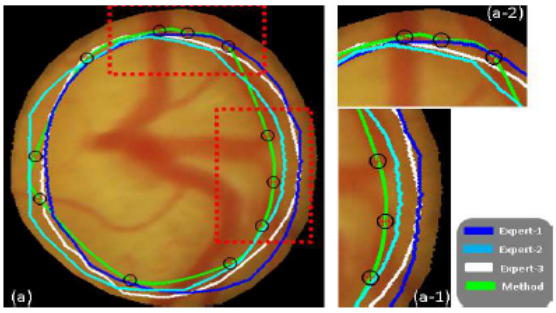


Fig. 5: Detected cup boundary.

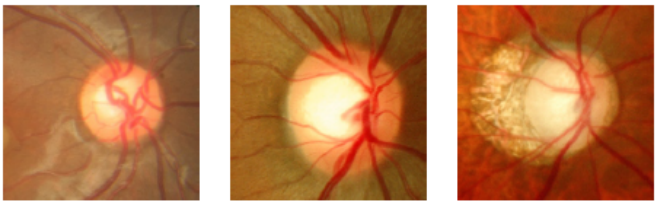
To select relevant vessel bends (called r-bends), a two-stage filtering process is used. In the first stage, bends close to the bright pallor region inside the OD are selected as coarse candidates. In the second stage, anatomical knowledge (sector-wise location and vessel orientation) is used to refine these r-bends further. The final set of r-bends typically includes bends formed by vertical vessels in superior and inferior sectors and horizontal vessels in temporal and nasal sectors. The optic cup boundary is then estimated by fitting a 2D cubic cardinal spline through the selected r-bends to ensure smoothness across sectors with sparse bend information.

The architecture offers good sensitivity in detecting cup-to-disc ratio (CDR) differences between normal and glaucomatous images. It shows high agreement with expert markings, especially for glaucomatous images with visible vessel bending. However, one key limitation is that the method struggles in regions where no 2D vessel kinks exist, especially in sectors with naturally low vessel density. Another limitation is that the method depends heavily on accurate vessel segmentation, making it sensitive to segmentation errors. Additionally, the ROS-based bend detection is scale-dependent and may misclassify noise or low-contrast vessel curves as bends. The approach also lacks integration of 3D depth information, limiting its performance in images where 2D features alone are insufficient to locate the true cup boundary. Finally, the method was evaluated on a relatively small dataset (133 images), limiting its generalizability across larger and more diverse datasets.

2.3 Glaucoformer: Dual-domain Global Transformer Network for Generalized Glaucoma Stage Classification

Das, Nayak, and Pachori 2025

Glucoformer is a deep learning-based model designed for glaucoma stage classification (Normal, Early, Advanced) from retinal fundus images without requiring prior segmentation.



(a) (b) (c)

Fig. 3. Example retinal fundus photographs: (a) normal eye, (b) early-stage, and (c) advanced-stage glaucoma.

The methodology starts by feeding the input fundus image into a DenseNet-121 backbone, which extracts crucial retinal features such as the brightness of the neuroretinal rim, the loss patterns of RNFL (retinal nerve fiber layer) and changes in blood vessels.

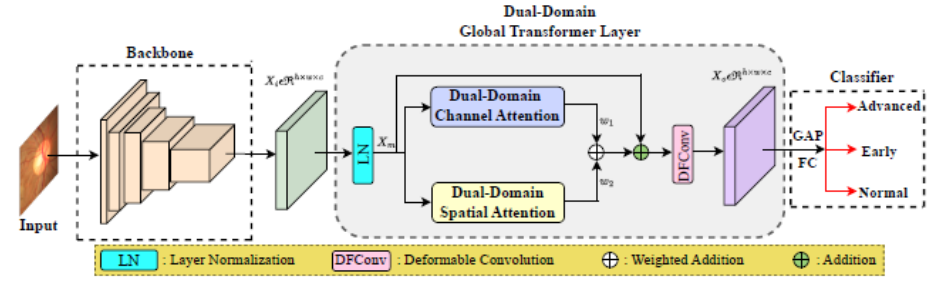


Fig. 1. Overall pipeline of our proposed Glaucoformer, comprising a backbone network followed by a dual-domain global transformer layer and a classifier, for the classification of glaucoma stages.

To capture the information in frequency and spatial domains, Glucoformer introduces a novel Dual-domain Global Transformer Layer (DGTL). This layer contains two main components: Dual-domain Channel Attention (DCA) and Dual-domain Spatial Attention (DSA). The DCA module helps the model identify and prioritize important feature channels across both spatial and frequency domains, while the DSA module learns spatial locations that need more focus, such as the optic disc and cup regions. Another unique component is the Fourier Domain Feature Analyzer (FDFA), which converts the features into the frequency domain using Fast Fourier Transform (FFT), processes them with convolutional layers, and then transforms them back to the spatial

domain using inverse FFT (iFFT). This process helps capture global shape, texture, and edge-related features like cupping and vessel boundaries. Additionally, deformable convolutions are integrated to allow the model to adapt to non-circular optic cups and curved blood vessels. Finally, a softmax classification layer predicts the glaucoma stage into one of three categories: Normal, Early, or Advanced. Overall, Glucoformer's methodology effectively combines CNN feature extraction, transformer-based attention, and frequency domain analysis for robust glaucoma stage detection.A key limitation of the Glucoformer model is the absence of clinical validation for tracking glaucoma progression over time.

TABLE III
COMPARISON OF GLAUCOFORMER WITH VIT FRAMEWORKS. 'BLUE' AND 'GREEN' COLOR FONTS INDICATE THE BEST AND SECOND-BEST RESULTS.

Method	A_{cc} (%)	F1 (%)	AUC	Params (M)	FLOPs (G)	Mem (MB)	Inf. T(ms)
ViT-b16 [23]	83.62	83.39	0.9419	86.57	35.24	330.24	19.39
ViT-b32 [23]	84.69	84.68	0.9368	87.46	8.84	333.62	25.38
SwinV2 [41]	83.83	84.03	0.9418	48.96	19.41	186.77	14.39
MaxViT [42]	84.48	84.47	0.9464	68.16	23.49	260.01	28.89
FlexiViT [43]	81.25	81.38	0.9270	21.68	10.74	82.69	18.45
CAFormer [44]	85.77	85.45	0.9492	24.29	8.29	92.69	17.07
Glaucoformer	**86.85**	**85.50**	**0.9505**	**9.85**	**6.02**	**37.56**	**13.09**

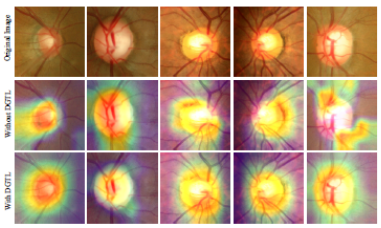


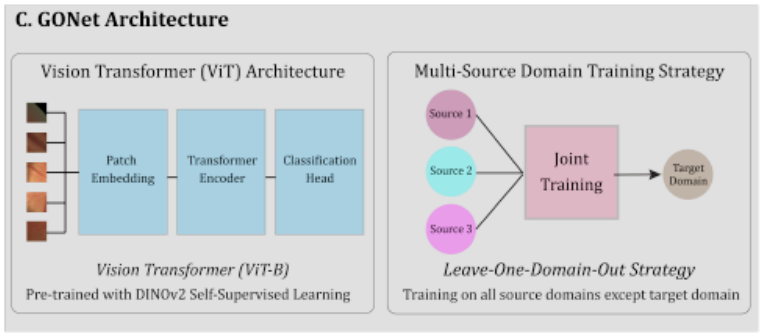
Fig. 5. Illustration of saliency maps generated using Grad-CAM++ for sample images from the HDV1 dataset. The first row depicts the original images, while the second and third rows demonstrate the saliency maps of DenseNet-121 and Glaucoformer, respectively.

2.4 GONet: A Generalizable Deep Learning Model for Glaucoma Detection

Abramovich et al. 2025

GONet is a deep learning-based model developed for the detection of glaucomatous optic neuropathy (GON) from color fundus photographs (CFPs). It utilizes a vision transformer architecture called DINOv2, which is pre-trained using self-supervised learning on a large set of natural images. The GONet pipeline begins with a quality check of the input fundus images using a model named FundusQ-Net to filter out low-quality images. Following this, the optic disc and cup regions are segmented using LUNet, which helps in calculating important clinical parameters like the Cup-to-Disc Ratio (CDR) and Rim-to-Disc Ratio (RDR). These segmented regions and image quality-checked inputs are then processed by the DINOv2 vision transformer for feature extraction and glaucoma classification. GONet was trained on a diverse set of over 119,000 images from seven different datasets, making it highly generalizable across different populations and imaging devices. Its training strategy involves multi-source domain training

to reduce domain bias and enhance robustness. The model achieved excellent performance, with AUC values ranging from 0.88 to 0.99 on unseen test domains. Unlike traditional models that rely heavily on CDR alone, GONet captures broader retinal features like rim thickness and blood vessel patterns for improved glaucoma detection. Overall, GONet represents a scalable and clinically applicable solution for automated glaucoma screening in real-world settings.



One major limitation of GONet is its lack of capability to track glaucoma progression over time, as it performs only binary classification (glaucoma or non-glaucoma) based on single static images. Additionally, GONet does not incorporate multi-modal clinical data such as Optical Coherence Tomography (OCT) scans, Intraocular Pressure (IOP) measurements, or patient visual field test results.

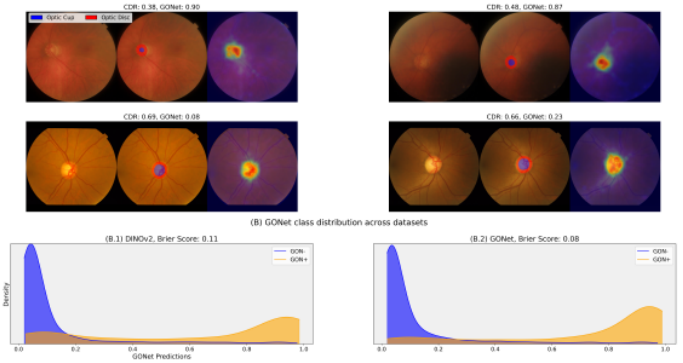


Fig. 3: Panel A: Examples of CFPs with low CDR (≤0.5) that are GON+ and CFPs with high CDR (≥0.65) that are GON−, and which were identified as such with certainty of ≥0.75. For each example, three images are shown: the original CFP (left), the segmentation highlighting the optic disc in red and optic cup in blue (center), and the model's attention map (right). Panel B: Distribution of DINOv2 (B.1) and GONet (B.2) predictions for GON+ and GON− CFPs over all target domains, using kernel density estimation (KDE). Brier score [45] is reported for each model.

2.5 Graph Convolution-Based Multi-task Model

Lenka, Mayaluri, and Panda 2025

The model proposes a multi-task deep learning framework that performs both optic

disc (OD) and optic cup (OC) segmentation and glaucoma classification from fundus images. It uses MobileNetV2 as the encoder, making it lightweight and suitable for medical imaging datasets. The decoder incorporates Graph Convolutional Networks (GCN) that model spatial relationships between pixel regions, improving boundary segmentation accuracy. Attention modules focus the model on the optic nerve head region, filtering out irrelevant background areas. The model calculates neuroretinal rim (NRR) features and uses the ISNT rule for more accurate CDR-based glaucoma detection. It was trained and evaluated on datasets like ORIGA, REFUGE, and DRISHTI-GS. The model achieved Dice scores of 97.95% for OD and 96.11% for OC, with a glaucoma classification accuracy of 97.43%. This integrated approach helps improve both segmentation and classification performance. However, the model requires accurate annotations for both segmentation and classification tasks during training. Its limitation includes dependence on segmentation quality for accurate glaucoma detection.

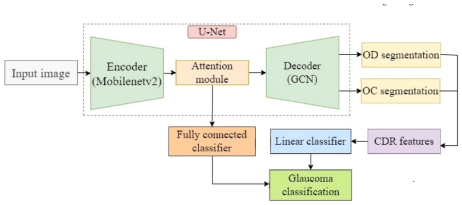
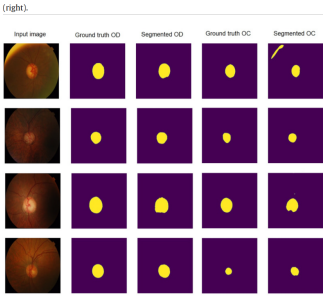


Fig. 2. Block diagram of proposed multi-task model.

Fig. 5. Cropped fundus images (left), cropped binary OD and OC segmented images (right).



2.6 A Hybrid Approach Combining EfficientNet and MRFO Optimization for Glaucoma Detection

Atia et al. 2025

This model combines EfficientNet-B0 for image feature extraction and a Feed-Forward Neural Network (FFNN) for processing numerical CDR values. The input includes both fundus images and pre-computed CDR values, making it a dual-input system. Preprocessing steps like Gaussian blur, gamma correction, and CLAHE improve image quality. Optic disc and cup segmentation are performed using K-Means clustering and morphological operations to estimate the CDR. The extracted features are concatenated and passed through fully connected layers with ReLU activation and softmax output for classification. MRFO (Marine Predators Algorithm) is used for hyperparameter optimization, tuning learning rate, batch size, and epochs. The model achieved very high classification accuracies, with over 99% on ORIGA and RIM-ONE DL, and even 100% on DRISHTI-GS dataset. It significantly outperforms other models like M-Net and CE-Net. However, the model doesn't use clinical data like IOP or visual field

tests. Its limitations also include lack of progression tracking and reliance only on static fundus images.

Algorithm 1: Pseudo-Code for Proposed Framework for Glaucoma Detection

Input: Fundus Images
Step 1: Image Preprocessing
For each image in dataset:
    Apply Gaussian Blur to reduce noise.
    Adjust brightness with Gamma Correction.
    Enhance contrast with CLAHE.
    Convert to grayscale for segmentation.
    Use K-Means to segment the image into regions.
    Create a binary mask of the optic disc.
    Refine mask using Morphological Operations and crop the optic disc
    Segment optic cup using K-Means on the cropped disc image.
    Refine the mask with Morphological Operations.
    Calculate diameters of the optic disc and cup using the largest contours.
    Compute CDR as given by Equation 3.
End for
Step 2: Store the preprocessed images and corresponding CDR values for further steps in the model.
Step 3: Feature Extraction:
For each preprocessed fundus image and its corresponding CDR value
    Pass the image through EfficientNet-B0 to extract deep image features.
    Pass the CDR value through the FFNN to extract relevant features.
    Concatenate the extracted image features from EfficientNet -B0 with the CDR features from the FFNN
End for
Step 4: Define a fully connected classification model for training, incorporating Dropout to reduce overfitting.
Step 5: Use the MRFO algorithm to optimize key hyperparameters, including learning rate, number of trainable layers, and number of epochs, optimizer selection, and batch size.
Step 6: Predict on the test dataset and calculate accuracy, precision, recall, and F1-score.
Output: Trained model, optimized hyperparameters, evaluation metrics, and predicted glaucoma labels.

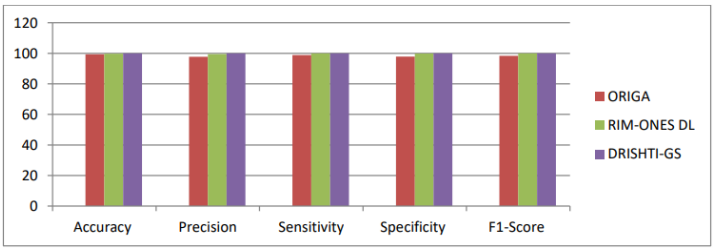


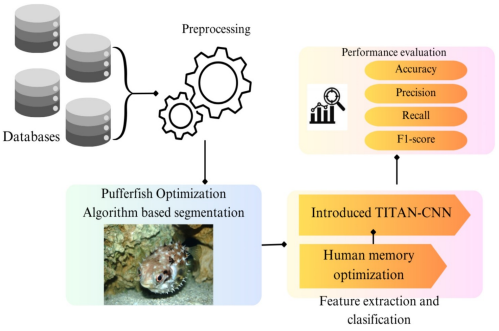
Fig. 8. Results of the proposed model across different datasets

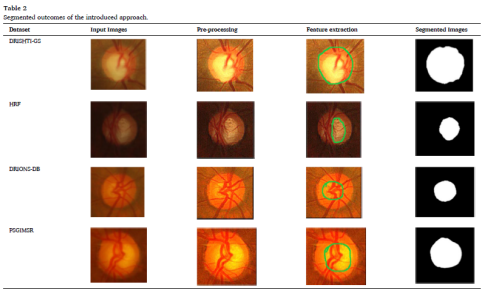
2.7 TITAN-CNN: Twin Inception Transformer Attention Network with Cycle-consistent CNN and POA segmentation model for glaucoma diagnosis

Rekha et al. 2025

TITAN-CNN, or Twin Inception Transformer Attention Network with Cycle-consistent CNN and POA segmentation, is a deep learning model developed for glaucoma detection from retinal fundus images. The architecture begins with a preprocessing step that uses adaptive morphological wavelet filtering and Perona-Malik diffusion to enhance the quality of input images. The POA (Pufferfish Optimization Algorithm) module performs precise segmentation of the optic disc (OD) and optic cup (OC) regions by mimicking the defense mechanism of a pufferfish to optimize the boundary regions. For feature extraction, TITAN-CNN employs the Twin-Inception Transformer, which

captures both local and global spatial features from multiple receptive fields. Following this, the Hinge Attention Network (HAN) emphasizes edge-level attention, focusing more on the cup-disc boundary regions where glaucoma-related changes are often visible. To ensure feature consistency across various transformations and image augmentations, the model integrates a Cycle-consistent CNN (Cycle-CNN). Furthermore, Human Memory Optimization Algorithm (HMOA) is used during training to enhance generalization and avoid overfitting.





The strengths of TITAN-CNN include its high segmentation accuracy, excellent AUC scores (up to 0.996), and robust performance across multiple public datasets like DRISHTI-GS and HRF. Its multi-module architecture effectively captures fine boundary details and global contextual features, making it suitable for accurate glaucoma diagnosis.

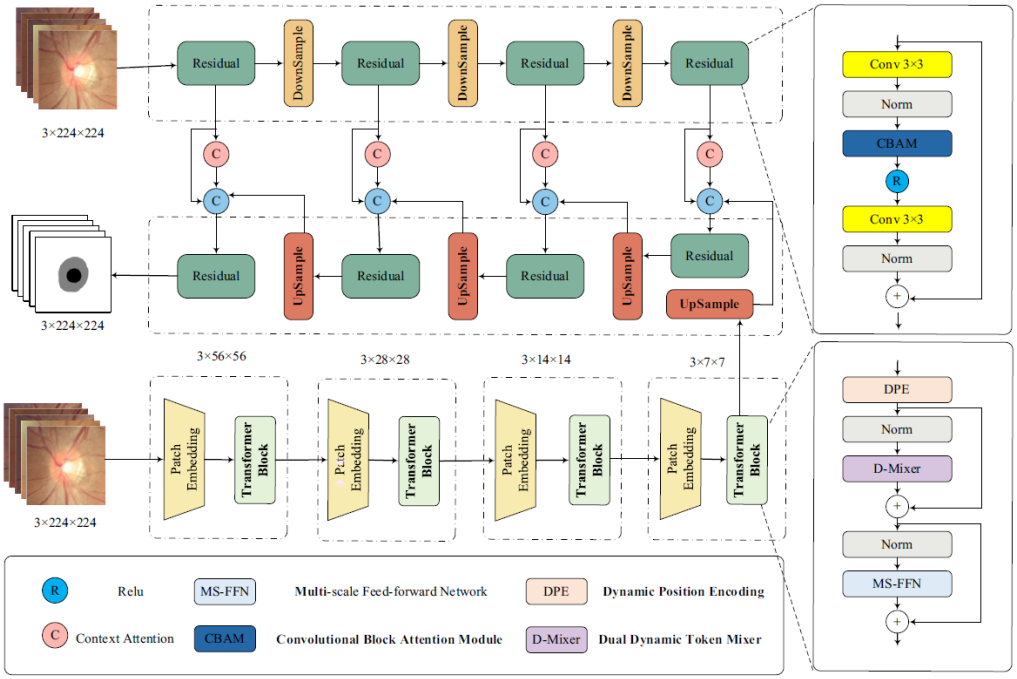
However, TITAN-CNN also has some limitations. The most critical drawback is its architectural complexity, making the model computationally expensive and memory-intensive, which limits its real-time deployment in clinical settings with low-resource hardware.

2.8 CC-TransXNet: Hybrid CNN-Transformer for OD/OC Segmentation

Yuan et al. 2025

CC-TransXNet focuses on accurate optic disc and cup segmentation from fundus images to improve glaucoma screening. The model uses a W-shaped encoder-decoder architecture that integrates CNN and Transformer blocks. Its CNN backbone, based on an improved ResNet, handles local feature extraction, while the Contextual Transformer (CoT) captures global contextual information. A CBAM (Convolutional Block Attention Module) adds spatial and channel-wise attention, helping the model focus on important regions. The dual decoder design fuses feature maps from both CNN and Transformer streams for precise segmentation. CC-TransXNet was tested on datasets

like REFUGE2, RIM-ONE DL, GAMMA, and Drishti-GS, achieving OD Dice scores up to 0.97 and OC Dice scores around 0.94. The model also delivered low CDR Mean Absolute Error (around 0.036 to 0.045). It consistently outperformed benchmark models like U-Net and Swin-UNet in both segmentation accuracy and generalization. However, the model has high computational demands due to its dual attention and dual decoder modules. It also lacks explainability features like saliency maps.



2.9 Boundary-Aware Transformer (BAT) for OD/OC Segmentation

Wang, Kim, and Eom 2025

The BAT model introduces a boundary-sensitive approach to segment the optic disc and cup with high precision in fundus images. It uses structure-preserving data augmentation based on truncated Gaussian sampling to generate diverse yet anatomically realistic training samples. The core innovation is the Boundary-Aware Transformer Attention (BAT) module, which enhances skip connections in a U-Net-like architecture by focusing on the inclusion relationship between the optic cup and disc. A geometry-aware loss function, combining Dice loss with normalized Hausdorff distance, ensures better boundary detection, especially for low-contrast regions. The model was evaluated on datasets like REFUGE, Drishti-GS, ORIGA, and G1020, showing superior segmentation accuracy compared to U-Net and Swin-UNet. The BAT module

helped significantly improve performance on boundary regions, where most CNN-based models struggle. Inference time was also fast (around 22 ms), making it suitable for real-time use. The model achieved high Dice scores, up to 0.98 for OD segmentation on REFUGE. However, BAT's performance under poor-quality images or rare anomalies remains uncertain. It also lacks external clinical validation on private datasets.

BAT module substantially enhances fine local boundary refinement around the boundaries.

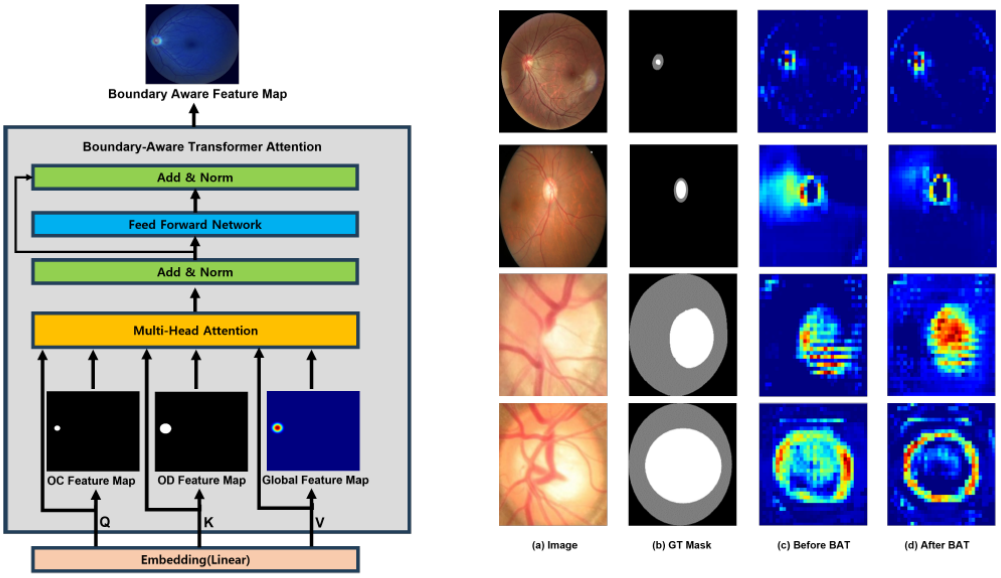
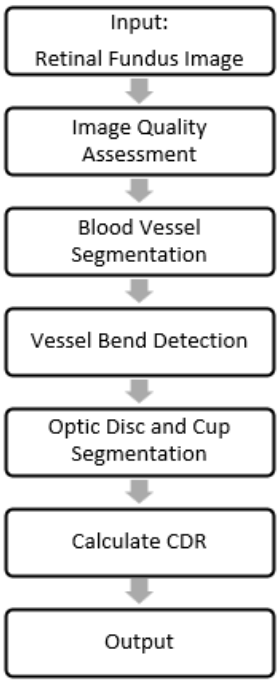


Figure 4. Structure of the proposed BAT module (**left**) and attention map visualization (**right**). The inputs to the attention mechanism are set as Query from the optic cup features, Key from the optic disc features, and Value from the global feature map, enabling boundary-aware representation learning.

Chapter 3

Design

Based on the insights gained from existing glaucoma detection models, I propose a new hybrid architecture that combines vessel bend analysis with deep learning-based segmentation. The design focuses on accurately identifying the optic cup and disc boundaries by leveraging blood vessel bending patterns detected through threshold-based curvature analysis.



3.1 Input: Retinal Fundus Image

The first step of the proposed model involves acquiring a high-resolution Retinal Fundus Image (RFI). This image provides a detailed view of the optic disc, optic cup, and retinal blood vessels. It serves as the primary input for all subsequent processing stages in the glaucoma detection pipeline.

3.2 Image Quality Assessment

After image acquisition, the input fundus image undergoes quality assessment. This deep learning-based model evaluates image clarity, contrast, and overall quality. Low-quality images affected by blur, poor lighting, or artifacts are filtered out to ensure only diagnostically useful images proceed further.

3.3 Blood Vessel Segmentation

The next step involves segmenting the retinal blood vessels using the proposed deep learning-based vessel segmentation model(UNET).The proposed model is specifically designed to capture fine, low-contrast, and highly branched vessels with greater accuracy. The output of this stage is a binary vessel mask highlighting all visible blood vessels in the retina .

3.4 Vessel Bend Detection using Thresholding

Once the vessel segmentation is complete, the model performs vessel bend detection. This is achieved using a threshold-based curvature analysis algorithm, which identifies points where the vessels show sharp directional changes. These high-curvature points are clinically significant as they often correspond to the boundary of the optic cup.

3.5 Optic Disc and Cup Segmentation

Guided by the detected vessel bend points, the model proceeds to segment the optic disc and optic cup regions. By incorporating vessel bend information as anatomical landmarks, the segmentation process becomes more accurate, especially in cases where the optic cup boundary is faint or irregular.

3.6 CDR Calculation

After successful segmentation, the model calculates the Cup-to-Disc Ratio (CDR), which is a crucial clinical indicator for glaucoma diagnosis. The CDR is computed based on the vertical or area ratio between the segmented optic cup and optic disc regions.

3.7 Output

The final output includes the segmented optic disc and cup boundaries, the calculated CDR value, and optionally, a visual overlay showing vessel bends and segmentation boundaries on the original fundus image. This comprehensive output can assist ophthalmologists in glaucoma screening and decision-making.

Chapter 4

Implementation

This chapter examines the implementation of the project.

4.1 Input

For the implementation, I used Kaggle as the coding platform because it provides free GPU access and easy dataset management. I utilized Retina Blood Vessel Segmentation Dataset on Kaggle, which contains retinal fundus images along with annotated vessel masks.It contained 80 training data and 20 testing data.

4.2 Blood Vessel Segmentation

In this stage of the proposed methodology, the focus was on segmenting the blood vessels from retinal fundus images as a preparatory step for vessel kinking analysis. Blood vessel segmentation is a crucial process because **vessel kinking**, also known as **bayoneting**, is a clinically recognized indicator of glaucoma.

In glaucomatous eyes, the optic cup (the central depression within the optic disc) often enlarges due to progressive damage to the optic nerve. This enlargement alters the normal path of retinal blood vessels, causing them to **bend or kink sharply** as they cross the edge of the expanded cup.

This distinct vessel bending, referred to as **bayoneting**, serves as a valuable anatomical cue for glaucoma detection in retinal fundus images. The following points highlight the importance of vessel kinking in glaucoma diagnosis:

- **Optic Cup Enlargement:** Glaucoma leads to optic nerve damage, which often results in the enlargement of the optic cup.

- **Vessel Displacement:** As the cup enlarges, blood vessels are displaced and forced to bend or kink to navigate the altered optic disc structure.
- **Bayoneting Sign:** This sharp vessel bending pattern, known as bayoneting, is widely recognized by ophthalmologists as a hallmark of glaucomatous changes.
- **Automated Detection Potential:** Recent research has focused on developing automated algorithms for detecting vessel kinks in fundus images, supporting early glaucoma screening efforts.
- **Other Structural Indicators:** Along with vessel kinking, structural features such as the Cup-to-Disc Ratio (CDR), Rim-to-Disc Ratio (RDR), and optic disc hemorrhages are also evaluated for comprehensive glaucoma diagnosis.
- **Importance of Early Detection:** Detecting vessel kinking at an early stage is crucial, as timely diagnosis and treatment can help slow disease progression and prevent irreversible vision loss.

For vessel segmentation, a U-Net-based deep learning model was implemented using the Retina Blood Vessel Segmentation Dataset from Kaggle.

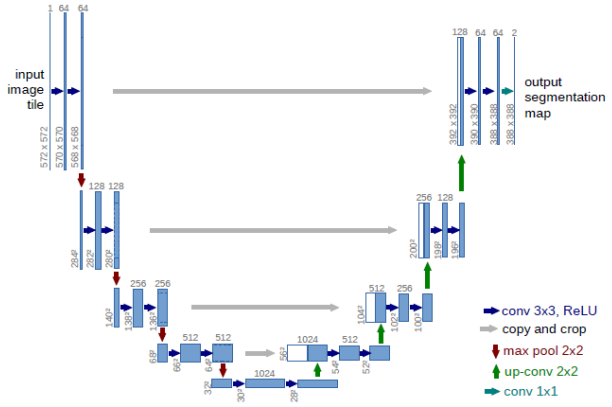


Fig. 1. U-net architecture (example for 32x32 pixels in the lowest resolution). Each blue box corresponds to a multi-channel feature map. The number of channels is denoted on top of the box. The x-y-size is provided at the lower left edge of the box. White boxes represent copied feature maps. The arrows denote the different operations.

Ronneberger, Fischer, and Brox 2015

The input retinal fundus images were resized to a resolution of 512×512 pixels to maintain consistency and reduce computational cost. The model consisted of a typical encoder-decoder structure with skip connections to retain spatial information across layers. The encoder path comprised four downsampling blocks, where each block contained two consecutive convolutional layers with 3×3 kernels, followed by batch normalization and ReLU activation functions, along with max-pooling layers for spatial

reduction. The number of feature channels doubled after each downsampling step, progressing from 64 to 512 channels.

At the bottleneck, the model utilized two 3×3 convolutional layers with 1024 filters, batch normalization, and ReLU activations to capture high-level abstract features. The decoder path used transposed convolution layers for upsampling, followed by two 3×3 convolutional layers with batch normalization and ReLU activations at each stage. Skip connections between corresponding encoder and decoder layers ensured the preservation of fine-grained spatial details. Finally, a 1×1 convolutional layer was used to produce the single-channel binary vessel segmentation mask.

The network was trained with a batch size of 2, using the Adam optimizer with a learning rate of 0.0005. To handle the class imbalance problem typically present in vessel segmentation tasks, a combined Dice and Binary Cross-Entropy (BCE) loss function was used as the objective criterion during training.The dataset provided both retinal fundus images and corresponding vessel masks. The dataset was divided into training and testing subsets. The model was trained using a **Dice-BCE Loss** function to address class imbalance effectively. The segmented vessel outputs were visually inspected and demonstrated satisfactory alignment with the provided ground truth vessel masks.

This blood vessel segmentation step serves as the foundation for the next phase of the research, which involves **detecting vessel kinking points** to assist in accurate optic cup boundary estimation.

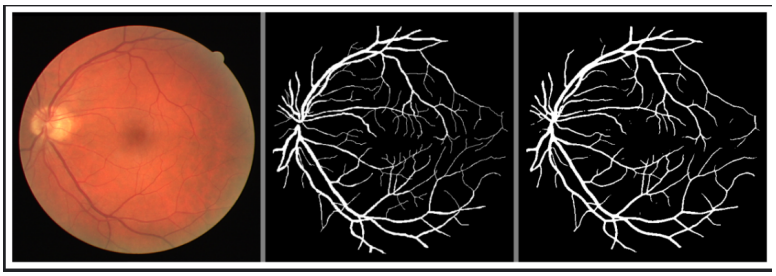
Chapter 5

Results

This chapter evaluates the overall project and provides results of tests carried out.

5.1 Segmented Blood Vessels

The performance of the implemented U-Net-based model for retinal blood vessel segmentation is illustrated in the figure below. The left panel shows the original retinal fundus image from the Kaggle dataset. The middle panel represents the ground truth binary vessel mask, which serves as the reference for evaluation. The right panel displays the predicted vessel segmentation output generated by the U-Net model. Visually, the predicted segmentation aligns well with the ground truth, successfully capturing both the major and medium-sized blood vessels. Although many thin vessels are also detected, there are still some instances where very fine vessels remain incomplete or missing, suggesting scope for further improvement. Overall, the model demonstrates promising performance in delineating vessel structures from the retinal images.



5.2 Result Analysis

The U-Net model was trained for vessel segmentation on retinal fundus images using a combined Dice and Binary Cross-Entropy (BCE) loss function. After training, the

model achieved a training loss of 0.3199 and a validation loss of 0.3463, indicating that the model was able to generalize reasonably well from the training set to the unseen validation set without significant overfitting.

To quantitatively evaluate the segmentation performance, Dice Coefficients (Dice Scores) were calculated separately for both training and validation datasets. The model achieved a Train Dice Score of 0.6801 (68.01%) and a Validation Dice Score of 0.6533 (65.33%), demonstrating a good overlap between the predicted vessel masks and the ground truth annotations on both datasets. This shows that the model can effectively segment blood vessels while maintaining consistency across different data splits.

Further evaluation was conducted using multiple standard classification metrics on the test set. The Jaccard Index (IoU) was 0.6836 (68.36%), indicating a high degree of intersection between predicted and true vessel regions. The F1-score, which balances precision and recall, reached 0.8115 (81.15%), showing that the model effectively balances false positives and false negatives. The recall was 0.8171 (81.71%), reflecting the model's ability to correctly identify most vessel pixels. The precision was 0.8138 (81.38%), indicating the accuracy of the predicted vessel regions. Additionally, the overall pixel-wise accuracy was high at 96.72%, meaning that most pixels across the test images were classified correctly as either vessel or background.

Finally, the Frames Per Second (FPS) metric was recorded as 221.77 FPS, highlighting that the model achieves fast inference speed, making it suitable for real-time or near-real-time applications in clinical settings.

Metric	Value
Train Loss	0.3199
Validation Loss	0.3463
Train Dice Score	0.6801
Validation Dice Score	0.6533
Jaccard Index (IoU)	0.6836
F1 Score	0.8115
Recall	0.8171
Precision	0.8138
Accuracy	96.72%
Inference Speed (FPS)	221.77

Chapter 6

Conclusion

In this research project, a U-Net-based deep learning model was successfully implemented for the task of segmenting retinal blood vessels using the publicly available Kaggle Retina Blood Vessel Segmentation dataset. The model was trained using a combined Dice and Binary Cross-Entropy (Dice and BCE) loss function to effectively balance both region-based accuracy and pixel-level precision.The experimental results confirm that the U-Net architecture, combined with appropriate loss functions and optimization techniques, can effectively segment retinal blood vessels with good accuracy and speed. However, minor performance gaps between training and validation suggest the need for further improvements in handling small vessel structures and reducing false positives. Future work will focus on incorporating vessel bend angle analysis, optic disc boundary detection using thresholding, and experimenting with advanced architectures like Attention U-Net or transformer-based models to further enhance segmentation quality and clinical applicability.

6.1 Future Work

Although the implemented U-Net-based model achieved satisfactory performance in retinal vessel segmentation, there are several areas for further improvement and extension:

Advanced Data Augmentation Techniques: Applying more diverse augmentation strategies such as elastic deformations, noise addition, and brightness/contrast adjustments to enhance model generalization.

Incorporating Attention Mechanisms: Implementing models like Attention U-Net to enhance focus on relevant vessel regions and suppress background noise.

Experimentation with U-Net Variants: Exploring advanced architectures such as U-Net++ and Residual U-Net for better feature representation.

Using Pretrained Encoders (Transfer Learning): Employing encoders like ResNet, VGG, or EfficientNet pretrained on large datasets to improve performance.

Learning Rate Scheduling and Optimizer Enhancements: Using learning rate schedulers like ReduceLROnPlateau or Cosine Annealing, and trying different optimizers like SGD with momentum.

Post-processing for Mask Refinement: Applying morphological operations and Conditional Random Fields (CRF) to improve mask sharpness and reduce false positives.

Training on Larger and Diverse Datasets: Incorporating public datasets like SMDG for better model generalization.

Deep Supervision: Introducing auxiliary outputs at intermediate layers to improve training convergence in deeper U-Net models.

Vessel Bend Angle Detection at the Optic Disc Boundary: Future work will also focus on analyzing vessel bending angles near the optic disc boundary, which is clinically relevant for glaucoma detection. This will involve iterating along vessel paths, threshold-based optic disc boundary extraction, to calculate the angles of vessel bends for diagnostic interpretation.

References

Abramovich, O. et al. (2025). "GONet: A Generalizable Deep Learning Model for Glaucoma Detection". In: *IEEE Transactions on Biomedical Engineering*. DOI: 10.1109/TBME.2025.3576688.

Atia, A. et al. (2025). "A Hybrid Approach Combining EfficientNet and MRFO Optimization for Glaucoma Detection from Fundus Images and CDR Values". In: *International Journal of Computers and Information* 12.2, pp. 17–29.

Das, D., D. R. Nayak, and R. B. Pachori (2025). "Glaucoformer: Dual-domain Global Transformer Network for Generalized Glaucoma Stage Classification". In: *IEEE Journal of Biomedical and Health Informatics* XX.X, pp. 1–12. DOI: 10.1109/JBHI.2025.xxxxx.

Joshi, G., J. Sivaswamy, and S. Krishnadas (2010). "Vessel Bend-Based Cup Segmentation in Retinal Images". In: *2010 20th International Conference on Pattern Recognition*. IEEE, pp. 2824–2827. DOI: 10.1109/ICPR.2010.690.

Lenka, S., Z. L. Mayaluri, and G. Panda (2025). "Glaucoma detection from retinal fundus images using graph convolution based multi-task model". In: *e-Prime - Advances in Electrical Engineering, Electronics and Energy* 11, p. 100931.

Rekha, R. et al. (2025). "TITAN-CNN: Twin Inception Transformer Attention Network with Cycle-Consistent CNN and POA Segmentation Model for Glaucoma Diagnosis". In: *Biomedical Signal Processing and Control* 85, p. 104042. DOI: 10.1016/j.bspc.2024.104042.

Ronneberger, O., P. Fischer, and T. Brox (2015). "U-Net: Convolutional Networks for Biomedical Image Segmentation". In: *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*. Springer, pp. 234–241. DOI: 10.1007/978-3-319-24574-4_28.

Wang, S., B. Kim, and D.-S. Eom (2025). "Boundary-Aware Transformer for Optic Cup and Disc Segmentation in Fundus Images". In: *Applied Sciences* 15.9, p. 5165. DOI: 10.3390/app15095165. URL: https://www.mdpi.com/2076-3417/15/9/5165.

Wong, D. W. K. et al. (2009). "Automated detection of kinks from blood vessels for optic cup segmentation in retinal images". In: *Proceedings of SPIE - The International Society for Optical Engineering*. Vol. 7260. SPIE, 72601J. DOI: 10.1117/12.810784.

Yuan, Z. et al. (2025). "CC-TransXNet: a hybrid CNN-Transformer network for automatic segmentation of optic cup and optic disk from fundus images". In: *Medical & Biological Engineering & Computing* 63. Received July 3, 2024 · Accepted Nov 7, 2024 · Published online Nov 27, 2024, pp. 1027–1044. DOI: 10.1007/s11517-024-03244-3.