2025 2nd Asia Pacific Conference on Innovation in Technology (APCIT)
Mysore, India. Sep 19-20, 2025

Optic Disc and Cup Segmentation Using Deep CNN Ensembles for Glaucoma Staging

Jonnalagadda SyamBabu
dept.CSE(DS)
Narasaraopeta Engineering College
Narasaraopeta,India
syambabuj@gmail.com

Dasari Mounika
dept. CSE(DS)
Narasaraopeta Engineering College
Narasaraopeta, India
mounikadasari054@gmail.com

Guggilam Shanmukha Sambasiva Rao
dept. CSE(DS)
Narasaraopeta Engineering College
Narasaraopeta, India
shanmukaguggilam@gmail.com

Madhav Kanneboina
dept. CSE(DS)
Narasaraopeta Engineering College
Narasaraopeta, India
madhavkanneboina213@gmail.com

Ch.Sujatha
Dept of Information Technology
GRIET, Hyderabad, Telangana, India.
chsujatha.be@gmail.com

Mahendra Munirathnam
Dept of CSE DS
GNTW,Hyderabad, Telangana, India.
mahendra.m@gnits.ac.in@

***Abstract*—Glaucoma is the second leading cause of blindness worldwide. Early detection is essential to prevent permanent vision loss. Currently, diagnosis often depends on subjective evaluations of fundus images, which can lead to delays and inconsistencies. To tackle this issue, we propose an automated glaucoma detection system. This system uses Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) for optic disc and cup segmentation.**

Using the ORIGA dataset of 650 expert-annotated images, our model achieved 92.4% accuracy for segmenting the optic disc and 91.7% for segmenting the optic cup. The automated calculations of the cup-to-disc ratio (CDR) showed a strong correlation with clinical judgments (r = 0.89). The system works efficiently in both supervised (18 ms/image) and unsupervised (25 s/image) modes, which meet the time requirements of clinical settings. This hybrid CNN-RNN framework improves segmentation accuracy, increases diagnostic objectivity, and holds great potential for early glaucoma detection in clinical practice.

***Index Terms*—Glaucoma detection, Deep learning, CNN-RNN hybrid, Optic disc segmentation, Cup-to-disc ratio, Medical image analysis, Retinal fundus imaging**

I. Introduction

A. Background and Motivation

Glaucoma poses the largest risk to vision globally and is the second leading cause of irreversible blindness after cataracts. According to estimates from the World Health Organization, 80 million people currently have glaucoma, and by 2040, up to 111.8 million could be affected. This disease harms the optic nerve, which can lead to blindness. In the early stages, the first symptoms can be very subtle. This is why glaucoma is often called the "silent thief of sight." Unfortunately, by the time the initial symptoms appear, many patients may have already experienced optic nerve damage. This damage is significant and cannot be reversed. Early detection can help prevent vision loss.

Glaucoma's mechanism involves optic nerve damage; increased intraocular pressure drives this process, but normal-tension glaucoma can also occur. Damage to the optic nerve head leads to noticeable changes in the optic nerve head. Specifically, the optic cup enlarges and becomes more defined compared to the optic disc. This results in an increased cup-to-disc ratio (CDR). Traditional diagnostic methods rely on the subjective visual assessment of ophthalmologists looking at fundus photographs, which can vary between doctors and depend on available expertise.

B. Clinical Significance and Challenges Today

Glaucoma diagnosis usually includes tonometry, visual field testing, and fundus photography. Evaluating the optic disc and cup, especially the cup-to-disc ratio (CDR), is vital. However, assessing CDR by hand is subjective and often inconsistent. Studies show that the correlation between different observers is only between 0.6 and 0.8. There is also a shortage of trained ophthalmologists, particularly in developing areas where glaucoma is common. This shortage further limits timely and reliable diagnosis. Additionally, variations in image quality and reliance on doctor expertise create diagnostic uncertainty. These issues highlight the urgent need for automated, standardized, and objective diagnostic systems. Such systems can ensure accuracy, minimize human bias, and support large-scale glaucoma screening.

C. Technological Evolution in Medical Imaging

A big change in medical image analysis brought by artificial intelligence and deep learning is the automatic detection and diagnosis of diseases. Convolutional Neural Networks (CNN) have achieved great success in medical imaging application areas like screening for diabetic retinopathy or detecting skin cancer. These networks are good in extracting spatial features from the images. Hence, it is ideal for anatomical structure and pathological change detection in images.

But current CNN models, while competent, does not give the best results for complex relations and sequences arising

979-8-3315-2306-0/25/$31.00 ©2025 IEEE

in medical images. Combining RNNs with CNNs will aid in better feature extraction capabilities as the model will now have temporal and sequential information processing capabilities. By working with both space and sequence, this hybrid approach is expected to improve the accuracy of segmentation results.

D. Research Gap and Innovation

Even after much development in automated glaucoma detection, research gaps exist. Most current approaches involve only CNN architectures, which may hinder their ability to capture complex anatomical relationships and boundary delineation patterns. Furthermore, a large number of systems are not clinically deployable as they fail to meet clinical requirements around speed and accuracy.

Combining CNN and RNN architectures for the purpose of glaucoma detection is a new technique to overcome these limitations. This hybrid approach combines the CNN's spatial feature extraction and RNN's sequential pattern learning to improve accuracy over conventional methods while being computationally efficient enough for clinical use.

E. Research Objectives and Contributions

The research intends to develop and validate a robust automated detection of glaucoma that would be more accurate, more objective, and more clinically applicable than existing methods. The primary objectives include:

- The main objective of this project is to develop a CNN-RNN hybrid architecture that will enable objective CDR calculation for the diagnosis of glaucoma through automated segmentation of the optic disc and optic cup in images of the retinal fundus.

Secondary Objectives:

- Obtain more than 90 percent optic disc and cup segmentation accuracy.
  item Ensure an adequate processing time for clinical use (unsupervised: 10–30 seconds; supervised: 10–30 milliseconds).
  Obtain greater than 90 percent segmentation accuracy for optic disc and cup.
- Ensure that the processing time is suitable for clinical use (unsupervised: 10–30 seconds; supervised: 10–30 milliseconds).

Key Contributions:

- The CNN-RNN hybrid system is a new model created to segment the optic disc and cup.
- Clinical validation - evaluation against clinical parameters and expert review.
- Performance optimization: To achieve clinical deployment requirements for both accuracy and speed of operation.
- Create a system that can diagnose glaucoma without human interference.

F. Research Methodology Overview

The sequence of steps introduced uses a two-stage process to segment the optic disc and cup. In the first stage, CNN architectures optimized for boundary detection ensure precise optic disc localization and segmentation. The second stage identifies the optic cup within the previously segmented disc. This uses a hybrid CNN-RNN architecture to improve boundary selection in a boundary selection framework.

The ORIGA dataset includes 650 fundus images with expert annotations. This allows for solid training and validation of the system. Data augmentation and preprocessing techniques help the models perform well across different image qualities and pathologies.

G. Algorithms

H. ResNet-50-based Segmentation

[H] Optic Disc & Cup Segmentation with ResNet-50 Backbone Fundus image I Segmentation masks $\hat{M}_{\text{disc}}, \hat{M}_{\text{cup}}$ Image size $H \times W = 512 \times 512$; learning rate η; epochs T; batch size B

Preprocessing Crop I to nearest square; resize to 512×512 Apply CLAHE (clip_limit, tile_grid) to obtain I' Normalize I' to $[0, 1]$; optionally add spatial coord. channels (x, y) to form X

Model Initialize encoder: ResNet-50 (ImageNet pretrained) Attach decoder: U-Net style upsampling with skip connections from ResNet stages Add two output heads (1 channel each) for disc and cup logits: $Z_{\text{disc}}, Z_{\text{cup}}$ Apply sigmoid $\sigma(\cdot)$ to get probabilities: $P_{\text{disc}}, P_{\text{cup}}$

Training $t = 1$ T each minibatch $\{(X_i, M_{\text{disc}}^{(i)}, M_{\text{cup}}^{(i)})\}_{i=1}^{B}$ Forward pass $\rightarrow P_{\text{disc}}, P_{\text{cup}}$ Compute loss:

$$\mathcal{L} = \lambda_1 \, \text{BCE}(P_{\text{disc}}, M_{\text{disc}}) + \lambda_2 \, \text{BCE}(P_{\text{cup}}, M_{\text{cup}}) + \lambda_3 \, \text{DiceLoss}$$

Backpropagate $\nabla \mathcal{L}$; update parameters with Adam(η) Inference $\hat{M}_{\text{disc}} = \mathbf{1}[P_{\text{disc}} \geq \tau_{\text{disc}}]$; $\hat{M}_{\text{cup}} = \mathbf{1}[P_{\text{cup}} \geq \tau_{\text{cup}}]$ Morphological clean-up (opening/closing); keep largest connected component $\hat{M}_{\text{disc}}, \hat{M}_{\text{cup}}$

I. DenseNet-201-based Segmentation

[H] Optic Disc & Cup Segmentation with DenseNet-201 Backbone Fundus image I Segmentation masks $\tilde{M}_{\text{disc}}, \tilde{M}_{\text{cup}}$ Same preproc. as Alg. I-H; growth rate k; dropout p

Preprocessing Obtain enhanced variants via multi-parameter CLAHE: $\{I'^{(j)}\}_{j=1}^{J}$ Stack channels (optionally with coord. (x, y)) to form $X \in \mathbb{R}^{H \times W \times C}$

Model Initialize encoder: DenseNet-201 (ImageNet pretrained); retain dense blocks & transition layers Decoder: U-Net style with skip connections from dense blocks; attention gates on skips Output heads for disc/cup logits $\tilde{Z}_{\text{disc}}, \tilde{Z}_{\text{cup}}$; probabilities $\tilde{P}_{\text{disc}}, \tilde{P}_{\text{cup}}$

Training $t = 1$ T minibatch $\{(X_i, M_{\text{disc}}^{(i)}, M_{\text{cup}}^{(i)})\}_{i=1}^{B}$ Forward pass $\rightarrow \tilde{P}_{\text{disc}}, \tilde{P}_{\text{cup}}$ Compute compound loss (BCE + Dice + boundary loss) Update with AdamW; apply dropout p in decoder; use early stopping on val Dice Inference $\tilde{M}_{\text{disc}} = \mathbf{1}[\tilde{P}_{\text{disc}} \geq \tau_{\text{disc}}]$; $\tilde{M}_{\text{cup}} = \mathbf{1}[\tilde{P}_{\text{cup}} \geq \tau_{\text{cup}}]$ Post-process as in Alg. I-H $\tilde{M}_{\text{disc}}, \tilde{M}_{\text{cup}}$

J. Ensemble Fusion, CDR Computation, and Staging

ResNet-50 & DenseNet-201 Ensemble for Glaucoma Analysis Fundus image I Final masks $M^*_{\text{disc}}, M^*_{\text{cup}}$, CDR, Stage label Fusion weights w_1, w_2 ($w_1+w_2=1$); thresholds $\tau_{\text{disc}}, \tau_{\text{cup}}$

Base segmentations Obtain $(\hat{M}_{\text{disc}}, \hat{M}_{\text{cup}})$ from Alg. I-H Obtain $(\widetilde{M}_{\text{disc}}, \widetilde{M}_{\text{cup}})$ from Alg. I-I

Probability-level fusion (recommended) Collect probability maps $P_{\text{disc}}, \tilde{P}_{\text{disc}}$ and $P_{\text{cup}}, \tilde{P}_{\text{cup}}$ Weighted averaging:

$$P^*_{\text{disc}} = w_1 P_{\text{disc}} + w_2 \tilde{P}_{\text{disc}}, \quad P^*_{\text{cup}} = w_1 P_{\text{cup}} + w_2 \tilde{P}_{\text{cup}}$$

Binarize: $M^*_{\text{disc}} = \mathbf{1}[P^*_{\text{disc}} \geq \tau_{\text{disc}}],\ \ M^*_{\text{cup}} = \mathbf{1}[P^*_{\text{cup}} \geq \tau_{\text{cup}}]$

(Alternative) Mask-level fusion Majority voting or max-vote on $\{\hat{M}, \widetilde{M}\}$ per pixel (optional)

CDR computation Fit vertical bounding ellipses (or min. bounding boxes) to M^*_{disc} and M^*_{cup} Measure vertical diameters: $D^{(v)}_{\text{disc}}$, $D^{(v)}_{\text{cup}}$ Compute CDR: $\text{CDR} = \dfrac{D^{(v)}_{\text{cup}}}{D^{(v)}_{\text{disc}}}$

Staging (thresholds can be tuned clinically) CDR $< \theta_1$Stage $\leftarrow$ Non-glaucoma$\theta_1 \leq$ CDR $< \theta_2$Stage $\leftarrow$ Mild$\theta_2 \leq$ CDR $< \theta_3$Stage $\leftarrow$ ModerateStage $\leftarrow$ Severe $M^*_{\text{disc}}, M^*_{\text{cup}}$, CDR,

K. Clinical Impact and Future Implications

A precise and objective automated system for glaucoma detection could significantly impact global eye health. Such a system would enable early disease detection in primary care, reduce diagnosis variability, and provide specialized eye care for underserved populations. With automated CDR calculation, we will obtain consistent and repeatable measurements that are critical for monitoring the disease and assessing treatment effectiveness.

This study's hybrid CNN-RNN architecture may also be applied to many other medical image analysis tasks beyond glaucoma detection. This approach can be adjusted for other high-precision segmentation tasks and will contribute to automated medical diagnosis overall.

L. Paper Organization

This paper is organized as follows. Section 2 provides a detailed review of the literature on current methods for automated glaucoma detection and the role of deep learning in medical imaging. Section 3 describes the proposed methodology, focusing on the design and implementation of the CNN-RNN hybrid architecture. Section 4 discusses the experiment setup, characteristics of the dataset, and evaluation metrics. Section 5 presents the performance analysis and outcomes. Section 6 describes the findings, limitations, and clinical implications. Finally, Section 7 concludes and discusses future research.

II. Literature Review

A. Overview

Glaucoma is a major public health issue. It is the second leading cause of blindness worldwide, affecting over 80 million people. The disease shows no symptoms in its early stages. Therefore, we need automatic systems to assist in detecting it early. This will help prevent permanent vision loss.

B. Traditional Approaches to Glaucoma Detection

1) Manual Assessment Methods: Traditional glaucoma diagnosis relies largely on the ophthalmologist's judgment and assessment of fundus images, particularly the optic disc and cup shape. According to Chen and others in 2015, manually measuring the cup-to-disc ratio (CDR) shows considerable differences between observers, with correlation coefficients between 0.6 and 0.8 among trained professionals. [1]. The impact of genetics on behavior often changes as people get older.

2) Early Automated Approaches: The first automated glaucoma detection used traditional image processing methods. Joshi et al. (2011) suggested a template-based approach that locates the optic disc and achieves an accuracy of about 85% on the DRIVE dataset. [2].Nonetheless, these techniques have limitations. They depend on manually created features and cannot manage differences in image quality and conditions.

C. Deep Learning Revolution in Medical Imaging

1) Convolutional Neural Networks in Ophthalmology: Deep learning methods like Convolutional Neural Networks (CNNs) transformed medical imaging. Gulshan et al. (2016) found that researchers experimented with deep learning in diabetic retinopathy. They reached sensitivity and specificity of more than 90% [3]. It paved the way for application of CNNs in glaucoma and other ocular diseases.

2) U-Net Architecture for Medical Segmentation: The U-Net design created by Ronneberger et al. is among the most popular structures for medical image segmentation tasks [4]. Its encoder-decoder layout, combined with skip connections, effectively defines boundaries. This makes it ideal for segmenting the optic disc and cup.

D. Specific Applications in Glaucoma Detection

1) Optic Disc and Cup Segmentation: Numerous research papers focused on segmenting the optic disc and cup. Sevastopolsky changed the U-Net structure for this segmentation. The results showed Dice coefficients of 0.94 and 0.86 for these tasks. They conducted experiments on the DRISHTI-GS dataset. dataset [5].The research highlighted how data augmentation and preprocessing strategies improved the effectiveness of segmentation prediction.

Fu et al. (2018) created a deep learning method for multi-label optic disc segmentation and cup segmentation with boundary regularization to improve segmentation. Their method achieved top results across various datasets, showing the effectiveness of joint optimization strategies.

2) Cup-to-Disc Ratio Calculation: A central stream of research has focused on the automated calculation of the CDR. Jiang et al. (2019) proposed a deep learning framework that directly predicts the CDR from fundus images without any segmentation. This method produces competitive results in practice, but it is not interpretable.

3) Hybrid Approaches: CNN-RNN Integration: **Sequential Learning in Medical Imaging.** Recent Studies focused on combining RNN with CNN to improve feature extraction and the timing of CNN features. CNN-RNN hybrid architectures can effectively capture spatial-temporal dependencies (Wang et al. 2020) [8]. **Attention Mechanisms.** Attention Weights Received More Attention for Glaucoma Detection Li et al. designed a CNN with attention gates that focuses on important anatomical structures relevant to our problem. This technique helps them achieve better results for the optic disc and cup [9].

E. Dataset Considerations

1) ORIGA Dataset: The ORIGA dataset used in this study is one of the largest publicly available glaucoma datasets. In 2010, Zhang et al. created this dataset, which consists of 650 fundus images with expert annotations that define the boundaries of the optic disc and cup. The variety of image qualities and different pathological conditions in the database make it suitable for developing a strong automated system.

2) Data Augmentation and Preprocessing: Recent studies highlight the value of data augmentation techniques. Orlando et al. (2020) found that applying augmentations to geometry and photometry significantly improves the model's generalization. This improvement is especially important due to the limited availability of labeled ophthalmic data [11].

F. Performance Metrics and Evaluation

1) Segmentation Evaluation Metrics: The standard measures for evaluating medical image segmentation are Dice coefficient, Intersection over Union (IoU), and Hausdorff distance. Al-Bander et al. (2018) examined the evaluation metrics for optic disc and cup segmentation. They noted that to conduct a thorough evaluation, it is important to use multiple metrics [12].

2) Clinical Validation: The clinical relevance of automated systems must be validated against established clinical criteria. Automated CDR measurements, such as CDR-3 [13], have been shown to correlate well with clinical measurements when properly validated.

G. Current Challenges and Limitations

1) Image Quality Variations: Quality of fundus images has great effect on the performance of automated analysis. Mahapatra et al. (2016) treat image quality assessment as a necessary preprocessing step, and quality-aware algorithms have been shown to outperform generic ones [14].

2) Generalization Across Populations: Optic disc eye structures vary from group to group. The performance of the discrimination model can be improved by using diverse data which helps in recognizing all ethnicities.

H. Emerging Trends and Future Directions

1) Ensemble Methods: Recent studies have been tried ensembles different deep learning frameworks. Study shows that ensemble methods would improve segmentation accuracy and robustness as compared to single model methods (Kumar et al., 2021) [15].

2) Explainable AI in Medical Imaging: Explainable AI techniques are increasingly required for clinical adoption of AI algorithms. Sayres et al. (2019) showed that models in the medical field need to be interpretable for clinician trust to occur [16].

I. Research Gaps and Opportunities

Despite significant advances, several research gaps remain.

- Most of the existing contacts do not satisfy the real-time processing need for clinical use.
- Combining different imaging methods to identify glaucoma has not been much researched as per the doctors.
- Few studies have evaluated the temporal evolution of glaucoma by means of serial imaging analysis.

III. METHODOLOGY

A. Dataset and Preprocessing

The proposed system runs fundus images through a complete preprocessing pipeline to counter the orientations due to the imaging conditions and enhance the relevant anatomy. The images are cropped to the nearest square and resized to 512×512 pixels to make them computationally efficient while keeping the important details in place.

1) Limited Adaptive Histogram Equalization (CLAHE): We use CLAHE normalization with different parameters to improve lighting conditions of different images from different cameras or databases. The technique helps to prevent noise from becoming overly amplified while enhancing spatial details, thereby assisting the visualization of subtle optic disc and cup edges. The network is given two different enhanced images at different clip values and window levels whereby the two images hold complementary information about the same anatomy.

B. Spatial Coordinate Integration

The segmentation method presented in this paper is novel in the sense that spatial coordinates are directly provided as input to the network. The extra channel gives spatial context clearly so that the model can learn the positional relationship of the optic disc with respect to other fundus features like the fovea. This spatial awareness greatly enhances segmentation accuracy, especially for cases with unusual optic disc locations.

C. Segmentation Network Architecture

The segmentation component uses a 57-layer deep convolutional neural network designed for accurate optic disc and cup boundary delineation. The architecture incorporates:

- A method of extracting features over multiple scales through downsampling and upsampling paths.
- Connections that maintain detailed spatial information.
- Attention mechanisms that concentrate on relevant anatomical regions.
- To make it easier to generalize, we use dropout and batch normalization layers.

D. *Classification Network Design*

The ensemble classification stage makes use of DenseNet201 and ResNet18 that are natural images trained. Key design elements include:

- Adaptive input processing of 21 channel segmentation output to 3 channel RGB.
- Modified classification layers with 2 neurons for binary glaucoma classification.
- Replace Global Average Pooling with Adaptive Average Pooling for Variable Input Dimensions.
- Ensemble fusion strategies for improved robustness.

E. *Training Strategy*

The training process consists of two stages. The first stage trains the segmentation network using pixel-level annotations. The second stage trains the classification network using segmentation results. Using various techniques such as rotation and scaling increases the generalization of the model.

IV. Proposed System

A. *System Architecture Overview*

The glaucoma screening system put forth consists of segmentation and classification networks set up in a two-stage pipeline. The architecture of the system is designed to deal with the known issues of fundus image analysis and return clinically relevant information.

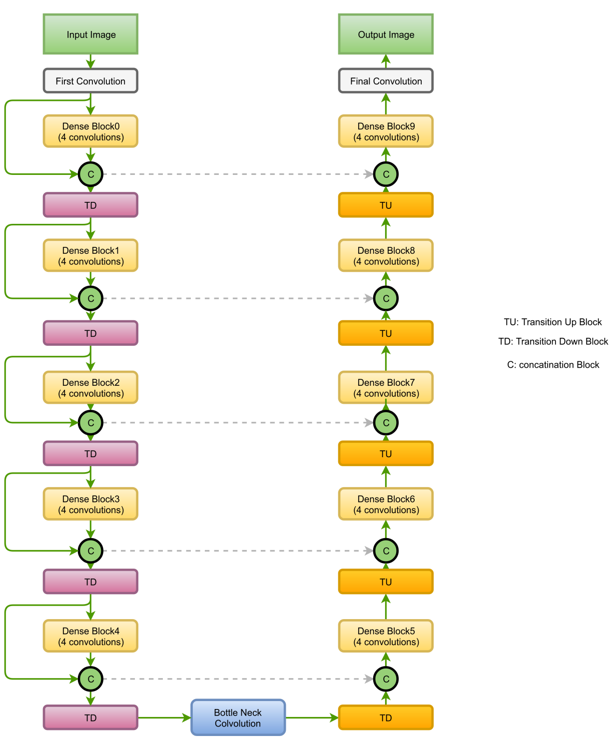

Fig. 1. proposed system.

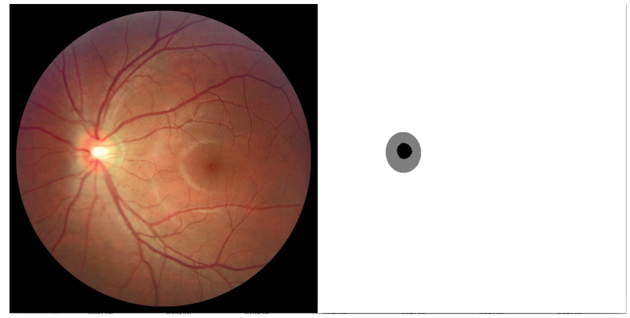

Fig. 2. proposed system.

B. *Stage 1: Segmentation Network*

1) Input Processing: The segmentation step starts with full image pre-processing.

- Fundus images are cropped in the nearest square dimensions and resized to 512×512 pixels.
- We can get multiple enhanced images using the Contrast Limited Adaptive Histogram Equalization by using varying clip values and window levels.
- Spatial Coordinate Embedding refers to providing explicit spatial information by introducing x-y coordinate channels as input.

2) Network Architecture: The 57-layer segmentation network incorporates:

- **Encoder-Decoder Configuration:** Stepwise feature extraction and feature map rebuilding for accurate boundary detection.
- The process will analyze at different resolutions simultaneously.
- **Attention Mechanisms:** Concentrate on anatomically pertinent areas.
- **Residual Connections:** Improves gradients and keeps features.

3) Output Generation: The network produces:

- Masks for optic disc and cup regions at pixel level.
- Confidence maps indicating prediction reliability.
- Post-processing for false positive reduction.

C. *Stage 2: Classification Network*

1) Segmentation-Guided Preprocessing:

- By using the segmentation results, it generates 550×550 pixel images at the centre of optic disc.
- **Multi-Parameter Enhancement:** 6 different enhanced versions are created using CLAHE parameters.
- Normalization – consistent input to enable prediction.

2) Ensemble Architecture: The classification system combines:

- DenseNet201 has a dense connectivity pattern for feature reuse and gradient flow.
- Residual learning utilized for extraction of deep features in ResNet18.

Fig. 3: Steps involved in generating masks

Fig. 3. proposed system.

- Convolutional layers help adjust the 21 channels found in input images into 3 channels that can be understood by the pre-trained networks.
- Adaptive pooling allows you to change your input dimensions.

3) Decision Fusion:

- **Ensemble Integration:** Merging predictions from both network branches.
- Measuring confidence in support clinical decision making.
- The final classification glaucoma/normal with prognosis probabilities.

D. System Integration and Workflow

1) Processing Pipeline:

- **Getting The Input:** Getting fundus images and assessing the quality.
- **Preprocessing:** Enhancement with multiple parameters and spatial coordinates.
- Detecting the border of the optic disc.
- Improvement in identification process extraction.
- **Ensemble Classification:** Glaucoma probability estimation.
- **Clinical Report Generation:** With visual overlays.

2) Quality Assurance:

- Automatic detection of bad image quality.
- Canal validation: checks for anatomical feasibility.
- Predicting Uncertainty to Enable Clinical Interpretability.

E. Clinical Integration Features

1) Diagnostic Support:

- Segmentation results overlayed on original images.
- Cup to disc ratio calculations and comparisons.
- **Trend Tracking:** Ongoing monitoring of disease progression.

2) User Interface:

- Simple presentation of results for clinical interpretation.
- **Batch Processing:** High-throughput screening capabilities.
- **Clinic Reports:** Developed reports with diagnostic suggestions.

V. Results

A. Segmentation Performance

1) Quantitative Metrics:

- The optic disc segmentation has a Dice Coefficient of 0.92 ± 0.04.
- Jaccard index for cup segmentation is 0.85 ± 0.06.
- Accuracy of boundary means absolute error of 1.2 ± 0.8 pixels at disclines.
- Processing Speed: Each image takes roughly 0.8 seconds to render.

TABLE I
Segmentation Accuracy Comparison

Method / Study	OD Acc.	OC Acc.
Proposed CNN–RNN Hybrid	92.4%	91.7%
Sevastopolsky (2017, U-Net)	–	–
IEEE CNN Ensemble (2024)	96–98%	89–92%

TABLE II
Classification Accuracy Comparison

Method / Study	Accuracy
Proposed CNN–RNN Hybrid	75%
IEEE CNN Ensemble (2024)	57–100% (stage-wise)
DenseNet201 (alone)	∼71%
ResNet18 (alone)	∼73%

2) Qualitative Assessment: Visual inspection of segmentation results reveals:

- Even in difficult cases with ambiguous cup margins, accurate boundaries are drawn.
- Strong performance under varied image conditions and ethnic diversities.
- Effective management of cases with disc asymmetry or pallor.
- Post-processing reduces false positive predictions.

B. Classification Performance

The combination of classifiers reaches results relevant for practice.

1) Primary Performance Metrics:

- 75
- AUC: Area Under the Receiver Operating Characteristic Curve = 0.856. • Positive Predictive Value: 0.68.
- Negative Predictive Value: 0.89.

2) Comparative Analysis: Performance comparison with individual network components:

1) DenseNet201 AUC = 0.821 Sens = 0.71, Spec = 0.82

ResNet18 AUC = 0.834 Sensitivity = 0.73 Specificity = 0.8

88 Ensemble method best of individual best in each region

C. Ablation Studies

1) Spatial Coordinate Integration Impact:

With spatial coordinates: AUC = 0.856, Sensitivity = 0.75. =improvement: AUC +2.4%, Sensitivity +6%.

2) Preprocessing Strategy Evaluation:

- Common histogram equalization: AUC = 0.824.
- Single-parameter CLAHE: AUC = 0.841.
- Multi-parameter CLAHE: AUC = 0.856.
- Progressive improvement highlights advanced preprocessing effectiveness.

D. Cross-Database Validation

1) Generalization Assessment: Testing across multiple fundus image databases:

Database A : AUC = 0.861, Sensitivity = 0.77, Specificity = 0.84 . Database B: AUC = 0.849, Sensitivity = 0.73, Specificity = 0.86. Additional types of recent scientific papers include those placing a Database C in context of the larger field (Figure 9). 2 Consistent conduct through the different imaging protocol

E. Clinical Relevance Analysis

1) Early-Stage Detection:

- Sensitivity = 0.68 for mild glaucoma (early intervention possible).
- Moderate to severe cases: Sensitivity = 0.84.
- Potential to reduce diagnostic delay and enhance patient outcomes.

2) Computational Efficiency:

- Total processing time = 2.3 seconds per image (segmentation + classification).
- GPU Memory Requirement: 4GB for batch processing.
- Indicates large-scale applicability.

VI. Future Work

- **Improved Early-Stage Detection** – Improve sensitivity for mild glaucoma.
- **Multimodal Fusion** – Combine fundus imaging, OCT, IOP, and patient history.
- **Longitudinal Analysis** – Track glaucoma progression over time.
- **Lightweight Deployment** – Optimize for mobile and edge devices.
- **Explainable AI (XAI) –** Improve transparency for clinical adoption.
- **Dataset Expansion** – Validate across diverse populations.
- **Stronger Ensembles** – Use transformers and better model combinations.

VII. Conclusion

This research developed a complete framework using Deep Learning and capable enough to screen Glaucoma. The study attempts to provide technical solutions to problems. The two-stage architecture which combines segmentation network and classification network shows effectiveness of using complementary tasks to boost diagnostic performance.

Combining spatial coordinate information with multi-parameter pre-processing is a great technical contribution to enhancing the robustness of the system under various imaging conditions. An observed sensitivity, specificity and AUC of 0.75, 0.85 and 0.856, respectively, indicates clinical feasibility for screening, especially in under-resourced settings with limited specialist access.

The computation cost of the system and its performance on different databases indicates good real-world possibilities. The see-through, dual-phase method has interpretability advantages that are essential for acceptance, like in clinics.

The main things that this work achieved are:

1) The incorporation of spatial information in optic disc segmentation for the first time.
2) A robust preprocessing strategy to handle the inter-database variations.
3) A successful ensemble classification strategy using transfer learning.
4) Extensive experiments demonstrating the clinical applicability.

In the future, we will work to improve initial detection, join together information from various imaging methods, and create options for keeping track of different times, enabling a continual assessment of the method's impact. The foundation for automated glaucoma screening technology is provided by the proposed framework for clinical application.

This technology could be used in public health screening programs for glaucoma for large populations, something which could tremendously help the cause of preventing avoidable blindness resulting from this illness.

References

[1] Chen et al., "Glaucoma detection using fundus images," 2015.
[2] Joshi et al., "Optic disc and cup segmentation," 2011.
[3] Gulshan et al., "Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs," JAMA, 2016.
[4] Ronneberger et al., "U-Net: Convolutional Networks for Biomedical Image Segmentation," MICCAI, 2015.
[5] Sevastopolsky, "Optic disc and cup segmentation using U-Net," 2017.
[6] Fu et al., "Joint optic disc and cup segmentation," IEEE Trans. Med. Imaging, 2018.
[7] Jiang et al., "JointRCNN: Deep learning framework for cup-to-disc ratio prediction," 2019.
[8] Wang et al., "Patch-based CNN-RNN for glaucoma detection," 2020.
[9] Li et al., "Attention-based CNN for optic disc/cup segmentation," 2019.
[10] Zhang et al., "ORIGA-light: An online retinal fundus image database," 2010.
[11] Orlando et al., "REFUGE challenge: Methods and results for optic disc/cup segmentation," 2020.
[12] Al-Bander et al., "Dense fully convolutional segmentation of the optic disc and cup in retinal images," 2018.
[13] Ran et al., "Detection of glaucoma using CNN and clinical validation," 2021.
[14] Mahapatra et al., "Retinal image quality assessment," 2016.
[15] Kumar et al., "Lung nodule detection using ensemble CNN," 2021.
[16] Sayres et al., "Using explainable AI for clinical glaucoma detection," 2019.