Applied Information System and Management (AISM)
Volume 8, (2) 2025, p. 195–202
P-ISSN: 2621-2536; E-ISSN: 2621-2544; DOI: 10.15408/aism.v8i2.46658
©2025. The Author(s). This is an open acces article under cc-by-sa

Automated Glaucoma Detection and Classification from Large-Scale Fundus Image Dataset Using YOLOv8 and CNN

Sheikh Aminul Islam[1], Humana Khan[2], Taslim Taher[3]

***Abstract*—Glaucoma is a major eye condition that slowly damages the optic nerve and remains one of the top causes of permanent blindness around the world. This study presents an automated framework for early detection and classification of glaucoma using artificial intelligence techniques applied to large-scale retinal fundus image dataset of over 17,000 images. The optic disc (OD) and optic cup (OC) were localized using YOLOv8. Following this, we conducted Region of Interest (ROI) extraction and contour masking to isolate the OD and highlight critical regions for further examination. We extracted essential features, such as the Cup-to-Disc Ratio (CDR), Vertical CDR (VCDR), neuroretinal rim (NRR) thinning, and compliance with the ISNT (Inferior > Superior > Nasal > Temporal) rule, resulting in a detailed tabular dataset. For classification, we applied ML and DL models. YOLOv8 demonstrated superior detection precision and CNN led the classification models with 87.13% accuracy. The proposed method offers a reliable, automated solution that can support large-scale glaucoma screening in clinical settings. This framework has the potential to assist ophthalmologists by improving the speed and accuracy of early glaucoma diagnosis, reducing the risk of vision impairment in affected patients.**

***Index Terms*—Glaucoma detection, fundus images, CNN, machine learning, deep learning, YOLOv8.**

I. Introduction

Glaucoma is a progressive eye disease that damages the optic nerve, often due to increased intraocular pressure, and is a leading cause of irreversible blindness worldwide [1]. Commonly referred to as the "silent thief of sight," glaucoma typically remains asymptomatic until significant vision loss has occurred [2]. The motivation behind this study is driven by the critical need for early detection to prevent irreversible vision impairment and improve patient quality of life.

Globally, glaucoma affected approximately 76 million individuals by 2020, with projections indicating a rise to 111.8 million by 2040 [3]. In 2010, glaucoma had already caused blindness in 8.4 million people, and this number is expected to increase to 22 million by 2040 [4]. In Bangladesh, a 2022 survey by the Bangladesh Glaucoma Society reported that 13% of individuals over 35 are affected, with 2 million confirmed and 5 million suspected cases. Among 12,000 participants, 3.2% were diagnosed with glaucoma, and the prevalence has surged by 20% over the past two decades [5]. These alarming statistics underscore the urgent need for effective glaucoma screening programs worldwide.

Clinically, glaucoma is evaluated through key structural changes in the OD and OC. Diagnostic indicators include the CDR, NRR thinning, rim-to-disc area ratio, disc diameter, and compliance with the ISNT rule [8]–[10]. The NRR, located between the OC and the outer margin of the OD, is a critical region where glaucomatous damage is most evident [9]. Additionally, vertical elongation of the optic cup and defects in the retinal nerve fiber layer are hallmark signs of disease progression [11]. Fundus imaging remains the primary modality for assessing these clinical markers, but manual analysis is time-consuming, subject to human error, and may lead to delayed diagnosis.

Traditional glaucoma detection methods rely on extensive clinical testing, which is expensive and inaccessible in low-resource settings. Moreover, manual evaluation of the CDR from fundus images is prone to inter-observer variability. To address these challenges, deep learning (DL) and object detection technologies have emerged as promising alternatives. CNN have demonstrated excellent performance in medical image classification, offering automated and scalable solutions [12]. Additionally, YOLO object detection models have achieved real-time detection capabilities in various image analysis tasks, making them highly relevant for glaucoma screening.

However, many prior studies either focused solely on classification without precise localization of the OD and OC or utilized older versions of object detection models, which lacked

Received: 14 May 2025; Revised: 27 May 2025; Accepted: 17 June 2025.
*Corresponding Author

[1*]Sheikh Aminul Islam, Department of Computer Science and Engineering, Ahsanullah University of Science and Technology, Dhaka, Bangladesh (e-mail: sheiikhaminul@gmail.com).

[2]Humana Khan, Department of Computer Science and Engineering, Ahsanullah University of Science and Technology, Dhaka, Bangladesh (e-mail: humanakhn@gmail.com).

[3]Taslim Taher, Department of Computer Science and Engineering, Ahsanullah University of Science and Technology, Dhaka, Bangladesh (e-mail: taslim.cse@aust.edu).

the accuracy of newer architectures. There remains a gap in integrating advanced detection techniques like YOLOv8 with CNN-based classification pipelines for comprehensive glaucoma diagnosis.

To bridge this gap, this study proposes an automated framework that leverages YOLOv8 for accurate OD and OC localization, followed by clinical feature extraction and CNN-based glaucoma classification. By incorporating clinically significant features such as the ISNT rule and NRR thinning, this research aims to improve diagnostic precision and provide a scalable, cost-effective solution for early glaucoma detection, especially in low-resource environments.

II. Related Work

A. CNN-Based Models for Glaucoma Detection

Deep learning, particularly CNNs, has driven significant advancements in automated glaucoma detection from fundus images. Amer Sallam et al. [1] fine-tuned pre-trained CNN architectures such as AlexNet, VGG, GoogleNet, and ResNet using the LAG dataset (3,758 fundus images), achieving an accuracy of 86.9% in detecting optic nerve damage. However, these models primarily focused on CDR-based features, potentially overlooking other structural indicators. Similarly, Silvia Ovreiu et al. [6] employed DenseNet121 on the RETINA, ACRIMA, and RIM-ONE datasets, achieving 95.6% accuracy in early glaucoma detection. Although effective, the approach relied heavily on high computational resources and lacked clinical feature integration. Alexandre Neto et al. [10] evaluated Xception, ResNet152V2, and Inception ResNetV2 on RIM-ONE r3, DRISHTI-GS, and REFUGE datasets, successfully identifying glaucoma indicators like optic nerve damage and disc cupping. Yet, these models did not incorporate domain-specific clinical rules such as the ISNT rule or asymmetry analysis.

B. Segmentation Techniques

Accurate segmentation of the OD and OC is crucial for reliable CDR computation. Mamta Juneja et al. [2] proposed G-Net for OD and OC segmentation on the DRISHTI-GS dataset, achieving 95.8% and 93% segmentation accuracies, respectively. However, the method's reliance on traditional segmentation techniques may limit adaptability to complex clinical scenarios. WangMin Liao et al. [8] introduced EAMNet, integrating ResNet with multi-layer average pooling for optic nerve head localization. Their model achieved a Dice coefficient of 0.9 and an AUC of 0.88 on the ORIGA dataset. While this method improved localization, it required substantial training data to maintain generalizability. Reference [16] enhanced segmentation accuracy using an attention-based CNN, achieving mean Dice scores of 0.989 for OC and 0.999 for OD, but the approach's computational complexity could challenge real-time deployment in low-resource settings.

C. Hybrid and Ensemble Approaches

Hybrid models have been employed to improve diagnostic accuracy by combining deep learning and traditional classifiers. Also, [13] combined CNN feature extraction with SVM, KNN, and RF classifiers on 1,500 fundus images from Edo Eye Hospital, achieving a high accuracy of 99%. Such ensemble methods improve generalization by leveraging diverse classifiers but often suffer from increased model complexity, making them less practical in resource-constrained settings. Meanwhile, [14] advanced this concept by integrating DeepLabv3+ with ensemble learning on 2,787 retinal images, achieving 99.7% segmentation accuracy and 99.53% classification accuracy. Despite their impressive results, these systems often prioritize performance over clinical interpretability and can be computationally demanding.

D. Clinical Feature Integration

Studies focusing on clinical rule-based integration have addressed some of these shortcomings. While [15] introduced a cost-effective screening method combining CDR analysis with the ISNT rule, achieving 92.9% accuracy while emphasizing neuroretinal rim width in 113 images from Mettapracharak Hospital. Although promising, the study's limited sample size restricts its generalizability. A.A. Alqarni employed attention-enhanced CNNs to improve segmentation accuracy, achieving mean Dice scores of 0.989 for the OC and 0.999 for the OD, but the lack of integrated classification limits its clinical utility. Other research [17] compared the IST and ISNT rules in assessing neuroretinal rim width across 1,856 eyes, concluding that the IST rule offered better specificity (91.2%) and sensitivity (89.7%). These clinically focused studies contribute valuable diagnostic insights but often lack comprehensive automated pipelines that combine segmentation, classification, and clinical evaluation.

Many existing studies rely heavily on vertical CDR as the primary diagnostic feature, often neglecting critical clinical indicators such as NRR thinning, the ISNT rule, and asymmetry between the eyes. Furthermore, most models are evaluated on limited datasets without considering variations in image quality, patient demographics, and clinical settings. The proposed study addresses these gaps by integrating YOLOv8 for accurate optic disc and cup detection and using CNN-based classification that incorporates both deep features and clinical diagnostic rules. This combined approach aims to improve robustness, generalizability, and diagnostic relevance for early glaucoma detection.

III. Method

A. Dataset

This study uses a subset of the SMDG-19 Glaucoma Dataset, which is publicly available on Kaggle in PyTorch format [18]. The dataset includes 17,242 retinal fundus images, which are arranged into training, validation, and testing sets. Each of these sets features images categorized into two different classifications:

- Class 1: Fundus images diagnosed with glaucoma.
- Class 0: Healthy fundus images.

Figure 1 provides an overview of the dataset. This dataset aims to identify the OD and OC in the images, both of which

Applied Information System and Management (AISM)
Volume 8, (2) 2025, p. 195–202
P-ISSN: 2621-2536; E-ISSN: 2621-2544; DOI: 10.15408/aism.v8i2.46658
©2025. The Author(s). This is an open acces article under cc-by-sa

are crucial for glaucoma evaluation. Following segmentation, various features are extracted.

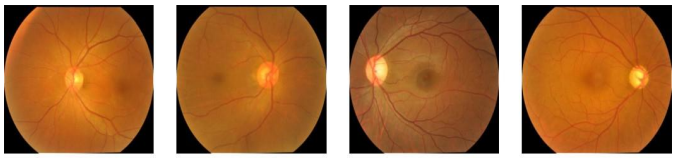

Fig. 1. Overview of the dataset, showcasing samples of the images used in this study.

To facilitate machine learning-based classification, these extracted features are organized into a structured tabular format. This tabular structure enables the implementation of a range of supervised learning models aimed at glaucoma detection and assessing its severity. The considerable size and variety of retinal images within the dataset enhance its robustness, making it an appropriate resource for developing and testing automated glaucoma detection systems.

B. *Proposed system*

This study proposes an automated and systematic approach for glaucoma detection from retinal fundus images, structured into three sequential stages: detection, feature extraction, and classification. The process begins with the detection of the OD and OC using a YOLO-based model, specifically trained to accurately localize these critical regions. Following detection, essential features such as the CDR, structural asymmetries, and other shape-based indicators are extracted to support reliable differentiation between glaucomatous and non-glaucomatous cases. In the final stage, ML and DL classifiers are employed to categorize the images based on the extracted features. This structured pipeline is designed to ensure smooth progression from detection to final classification, aiming to improve both diagnostic accuracy and processing efficiency. The complete workflow is illustrated in Fig. 2. By combining advanced object detection, feature engineering, and classification techniques, the proposed system enhances early glaucoma screening and supports timely clinical intervention.

1) *Detection,* to initiate the analysis, a YOLO-based model is employed to detect the OD and OC from dataset. This model is specifically trained to accurately locate these vital areas, which are crucial for further analysis and assessment of glaucoma.
2) *Feature extraction,* once the OD and OC are detected, key features—such as CDR, as well as shape and structural characteristics, are extracted. These features are essential for differentiating glaucomatous images from normal ones, providing critical data for classification.
3) *Classification*, in the final phase, the classification of retinal images is carried out to determine whether they are glaucomatous or normal. This classification stage utilizes ML and DL models which are trained on the extracted features. These models analyze the input data and provide the corresponding label based on their learned patterns and characteristics.

This approach follows a well-organized process, ensuring that each step flows smoothly into the final task of classification. The complete process of the framework is shown in Fig. 2, providing a clear visual representation of the workflow. This automated pipeline guarantees both accuracy and efficiency in the glaucoma screening process. By using advanced detection and classification methods, the approach helps with early diagnosis, which can greatly support clinical decisions and improve patient care.

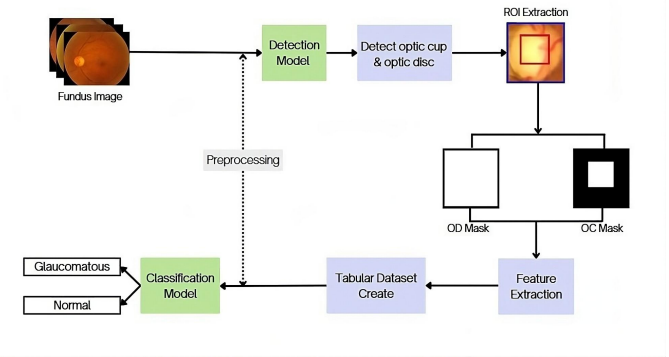

Fig. 2: Proposed system architecture.

C. *Dataset for Detection Task*

For the detection task, the dataset was carefully labeled to highlight the OD and OC regions, which play a key role in diagnosing glaucoma accurately. The images were annotated using the Computer Vision Annotation Tool (CVAT), a commonly used tool in computer vision projects for labeling and preparing data. Initially, polygons were drawn to outline the OD and OC areas, which were then converted into bounding boxes which are the required format of YOLO models.

Figure 3 illustrates an example of the annotated fundus images, showcasing the bounding boxes that indicate the OD and OC areas clearly. This process is vital for training the detection model, as it ensures the model can effectively locate and identify the important regions in retinal images.

D. *Data Preprocessing for Detection*

The dataset underwent several preprocessing steps. To maintain uniformity, all images were resized to 512×512 pixels and normalized before processing. Some data augmentation techniques were applied, such as flipping, adjusting brightness, rotating, and adding Gaussian noise. The bounding boxes were adjusted to match the augmented images accordingly. Furthermore, class balance was achieved to prevent bias during the training of the model.

E. *Detection of Optic Disc and Optic Cup*

In this study, YOLOv5 and YOLOv8 models were used to detect the OD and OC. These advanced object detection models were trained on an annotated dataset to accurately determine the OD and OC regions, which are crucial for glaucoma diagnosis. Both models were carefully evaluated to ensure

accurate localization, laying a strong foundation for the next steps of feature extraction and classification.

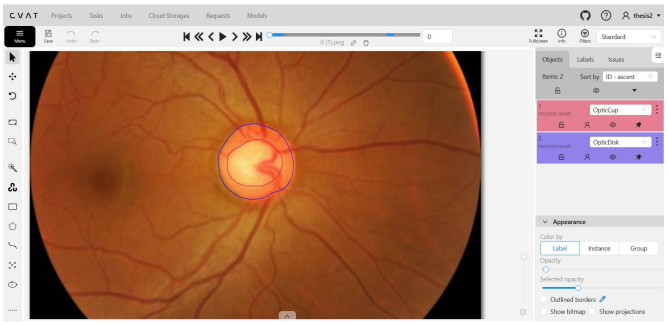

Fig. 3. Example of annotated fundus images using CVAT

YOLOv8 was chosen for its high speed, accuracy, and real-time processing capabilities, making it particularly well-suited for clinical applications. Its anchor-free architecture and advanced feature fusion significantly enhance the localization of small, overlapping structures, which is critical for accurate glaucoma screening. It provides fast and reliable detection with low computational cost, enabling scalable deployment even in resource-constrained clinical settings. Its performance in this study further validated its effectiveness and suitability for precise detection.

The results from the detection phase provided important bounding box coordinates for both the OD and OC, which enabled the extraction of key features necessary for classifying glaucoma. After this, the process moves on to isolating the Region of Interest, allowing the model to focus on the specific area for detailed analysis.

F. Extracting the Region of Interest (ROI)

In this research, the process of extracting the ROI begins with a detection model that identifies and marks the OD and OC in the input images. Once the detection phase is completed, the labeled results are utilized to localize these regions, enabling the extraction of the OD for further computation. The following step focuses on obtaining the ROI, which removes any extraneous areas and concentrates attention exclusively on the regions essential for the next stages of analysis.

To enhance focus on these important areas, the output image undergoes a process of ROI-based masking followed by contour masking. This methodical approach guarantees that only the pertinent features are retained for further processing, thereby improving the accuracy of later classification and analysis.

Figure 4 illustrates the entire procedure for ROI extraction, emphasizing the crucial steps of detection, isolation, and masking.

G. Feature Extraction

In the feature extraction stage, masks for the OD and OC derived from the ROI extraction were utilized. These masks made it easier to calculate several important features, such as the Vertical Cup-to-Disc Ratio (VCDR), Horizontal Cup-to-Disc Ratio (HCDR), Vertical Elongation, Neuroretinal Rim (NRR) Thinning, and adherence to the ISNT Rule. The detailed calculations of these features are presented below, emphasizing their importance in the overall analysis. Additionally, the theoretical concepts underlying each feature are examined to highlight their contributions to glaucoma diagnosis.

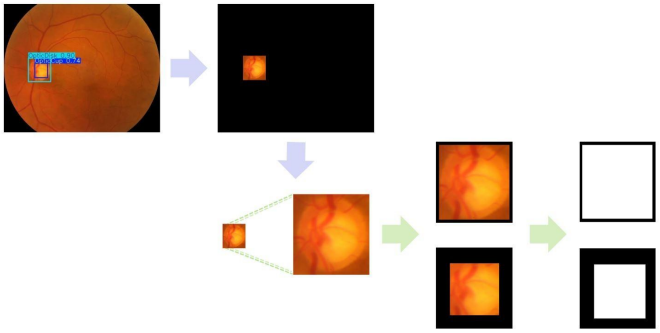

Fig. 4. ROI extraction procedure.

1) Vertical Cup-to-Disc Ratio (VCDR)

The VCDR plays a key role in glaucoma diagnosis, as it represents the proportion between the vertical diameter of the OC and that of the OD. OC is usually smaller. However, in glaucoma cases, it often expands vertically, resulting in an elevated VCDR. An elevated VCDR may suggest the presence of glaucomatous damage, as an increase in the vertical size of the optic cup often signals deterioration of the optic nerve. The VCDR can be calculated using the formula mentioned in (1).

$$\text{VCDR} = \frac{\text{Vertical Diameter of OC}}{\text{Vertical Diameter of OD}} \quad (1)$$

2) Horizontal Cup-to-Disc Ratio (HCDR):

HCDR is determined by taking the horizontal diameter of the OC and dividing it by the horizontal diameter of the OD. In some cases, the OC tends to expand horizontally, leading to a higher HCDR. An increased HCDR indicates a thinning of the NRR, suggesting possible harm to the retinal nerve fibers. The formula used to calculate the HCDR mentioned in (2).

$$\text{HCDR} = \frac{\text{Horizontal Diameter of OC}}{\text{Horizontal Diameter of OD}} \quad (2)$$

3) Vertical Elongation

Vertical elongation describes the abnormal enlargement of the OC in a vertical direction, resulting in a higher VCDR. In normal eyes, the OC retains a circular shape, but in cases of glaucoma, it appears more elongated. This vertical elongation is crucial for identifying glaucoma and evaluating its severity. It can be measured using a specific formula in (3).

$$\text{Vertical Elongation} = \frac{\text{VCDR} - \text{HCDR} + 1}{2} \quad (3)$$

4) Neuroretinal Rim (NRR) Thinning

NRR thinning is the decrease in thickness of the nerve

Applied Information System and Management (AISM)
Volume 8, (2) 2025, p. 195–202
P-ISSN: 2621-2536; E-ISSN: 2621-2544; DOI: 10.15408/aism.v8i2.46658
©2025. The Author(s). This is an open acces article under cc-by-sa

fibers within the NRR surrounding the OC. This thinning is a crucial sign of glaucoma, often beginning in the larger regions of the OD, which are especially vulnerable to injury. As glaucoma progresses, the thinning becomes increasingly evident and extends across the OD. The assessment of NRR thinning is expressed as follows in (4).

$$\text{NRR Thinning} = \frac{\text{Disc Diameter} - \text{Cup Diameter}}{2} \quad (4)$$

5) ISNT Rule Compliance

ISNT rule provides a framework for assessing the thickness of the NRR around the optic disc. According to this guideline, the expected thickness of the rim follows a specific sequence: Inferior > Superior > Nasal > Temporal. This indicates that the inferior segment of the optic nerve rim is usually the thickest, while the temporal segment is generally the thinnest. This pattern is typically seen in healthy individuals. However, in cases of glaucoma, this order can change, often resulting in greater thinning of the superior and inferior rims. The evaluation of adherence to the ISNT rule is conducted as follows in (5).

$$\text{ISNT (\%)} = \frac{\text{No. of ISNT checks passed}}{\text{Total ISNT checks}} \times 100 \quad (5)$$

H. Dataset for Classification Task

The extracted features were arranged in a table to facilitate glaucoma classification. Key features include the VCDR, HCDR, vertical elongation, thinning of the NRR, and compliance with the ISNT rule. These metrics are crucial for assessing the advancement of glaucoma. A snippet of the tabular dataset is presented in Table l.

Table 1.
Snippet of Tabular Dataset

Image	VCDR	HCDR	Vertical Elongation	NRR Thinning	ISNT	Labels
BEH-10.png	0.3905	0.3864	0.502	0.3026	1	0
BEH-109.png	0.466	0.4167	0.5247	0.2765	1	0
BEH-121.png	0.4111	0.4521	0.4795	0.2807	0.67	0
BEH-138.png	0.4471	0.7391	0.354	0.2009	1	0
BEH-149.png	0.4615	0.5303	0.4656	0.2486	0.33	1
BEH-156.png	0.5333	0.5506	0.4914	0.2265	0	1
BEH-158.png	0.5823	0.3763	0.603	0.2574	0.33	1
BEH-160.png	0.5289	0.5347	0.4971	0.232	1	0
BEH-165.png	0.5765	0.6265	0.475	0.1969	0	1
BEH-168.png	0.3176	0.2927	0.5124	0.3433	0.67	0
BEH-172.png	0.3468	0.3455	0.5006	0.3242	1	0
BEH-178.png	0.42	0.4845	0.4678	0.2711	1	0
BEH-183.png	0.7582	0.7473	0.5054	0.1223	0.67	1
BEH-189.png	0.5542	0.6081	0.473	0.2068	1	0

Several preprocessing procedures were implemented to prepare the dataset for classification, ensuring both data consistency and integrity.

I. Data Preprocessing for Classification

During the preprocessing phase, we removed unnecessary columns and addressed any missing values in the dataset. We then split the data, allocating 70% for training and 30% for validation purposes. SMOTE was utilized to help even out the class distribution by generating artificial samples for the smaller group.

J. Glaucoma Classification

To classify glaucoma, several algorithms from both machine learning (ML) and deep learning (DL) domains were applied. Additionally, CNN and VGG16 architectures were utilized by converting tabular data into image grids, making them suitable for convolutional models.

1) Machine learning classification models

Authors applied nine different machine learning models—Random Forest (RF), Logistic Regression (LR), Support Vector Machine (SVM), Decision Tree (DT), K-Nearest Neighbors (KNN), Naive Bayes (NB), XGBoost, LightGBM, and AdaBoost—to several tabular datasets. Despite conducting hyperparameter tuning, we observed only slight enhancements in the performance of the models. In Table 2, we outline the essential hyperparameters employed for training each model, which were fine-tuned to optimize classification performance.

Table 2.
Hyperparameters of ML Models

Model	Hyperparameters
LR	Penalty = l2; Solver = 'lbfgs'; Max iter = 500
SVM	Kernel = 'linear'; C = 10; Gamma = 0.1
KNN	N-Neighbors = 15; Metric = 'Manhattan'
RF	Estimators = 150; Criterion = 'gini'
DecisionTree	Criterion = 'gini'; Splitter = 'best'
XGBoost	Colsample bytree = 0.8; Gamma = 0.1; Lr = 0.05
AdaBoost	Estimators = 200; Lr = 0.01; Max depth = 3
LightGBM	Lr = 0.01; Estimators = 500; Subsamples = 0.7

2) Deep learning classification models

To capture detailed patterns in the data, deep learning architectures like MLP, TabNet, CNN and VGG16 were utilized. Additionally, CNN and VGG16 architectures were utilized by converting the tabular data into image grids to make it compatible with these models. A comparison was made between these DL approaches and traditional ML methods, highlighting the benefits of using DL for detecting glaucoma. CNN was chosen for its capability to effectively learn complex patterns from structured tabular data by capturing local feature interactions, enhancing classification performance.

- *Multi-layer perceptron (MLP)*
  The MLP architecture is designed for binary classification, starting with an input layer and multiple Dense layers with LeakyReLU activation. Techniques like Batch Normalization, L2 regularization, and Dropout enhance stability and reduce overfitting. The model ends with a sigmoid output layer for binary classification. MLPs are effective in capturing complex, nonlinear relationships, making them ideal for glaucoma detection with structured data.
- *TabNet*
  The TabNet model used in this study combines decision and attention networks to select important features and improve classification accuracy. It processes data over seven decision steps with an attention dimension of 128, using a relaxation factor of 1.5. The model was trained for 50 epochs with a batch size of 256, for binary classification.
- *CNN*
  The CNN model for glaucoma classification uses a Conv2D layer for feature extraction, followed by Flatten and Dense layers. A Dropout layer reduces overfitting, and a sigmoid output layer handles binary classification. Training uses the Adam optimizer and binary cross-entropy loss, with performance assessed via K-fold cross-validation, ROC curves, and confusion matrix.
- *VGG16*
  The VGG16 based model uses a pre-trained architecture for feature extraction in glaucoma classification. Tabular data is reshaped into 64×64×3 images to fit the model input. Once the top layers of VGG16 are removed, the extracted features are sent through a Flatten layer, then processed by a Dense layer with 128 neurons using ReLU activation, and a Dropout layer is applied to prevent overfitting. A sigmoid-activated output layer provides the binary classification. The model is trained using Adam and binary cross-entropy, and evaluated with accuracy, AUC, ROC, and confusion matrix.

While the method shows strong performance, it faces limitations such as dataset diversity, potential labeling errors, hardware constraints, and overfitting risks. Nonetheless, it advances prior work by integrating key clinical rules, combining image and clinical data features, and offering computational efficiency for real-time use in resource-limited settings.

IV. Results and Discussion

A. Detection Model Performance

The YOLOv8 model demonstrated superior detection performance compared to YOLOv5 across all evaluated metrics. As shown in Table 3, the detection performance of YOLOv8 and YOLOv5 was evaluated using precision, recall, and mean average precision (mAP) metrics. YOLOv8 outperformed YOLOv5 across all metrics, achieving a precision of 91.37% and a recall of 90.29%, indicating its superior accuracy and consistency in detecting the OD and OC.

Table 3.
YOLOv8 and YOLOv5 Performance

Model	Accuracy	
	YOLOv8	YOLOv5
RF	80.61%	79.1%
LR	81.7%	80.91%
SVM	82.1%	78.14%
DT	79.3%	75%
KNN	78%	76.13%
NB	81.45%	78.54%
XGBoost	81.24%	79.53%
LightGBM	81.53%	77.32%
AdaBoost	**83.22%**	80.28%

Additionally, YOLOv8 attained a higher mAP@0.5 of 92.13% and mAP@0.5:0.95 of 59.57%**,** demonstrating better localization performance across varying IoU thresholds compared to YOLOv5. This validates the selection of YOLOv8 for precise and reliable detection in this study. Figure 5 shows the confusion matrix for YOLOv8, highlighting its superior performance in detecting the OD and OC.

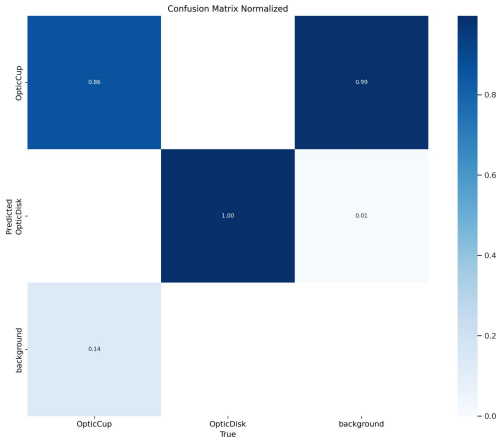

Fig. 5. Confusion matrix of YOLOv8 model.

B. Classification Model Performance

1) Machine learning model performance

In Table 4, classifiers using YOLOv8-extracted features consistently outperformed those using YOLOv5. AdaBoost achieved the highest accuracy (83.22%) with YOLOv8, followed by SVM (82.1%) and LR (81.7%). In contrast, YOLOv5 features yielded lower accuracies, with LR performing best at 80.91%. These results highlight YOLOv8's superior detection quality, leading to improved classification performance.

Table 4.
Performance of ML Models

Models	**Precision**	**Recall**	**mAP@0.5**	**mAP@0.5:0.95**
YOLOv8	0.91372	0.9029	0.9213	0.59573
YOLOv5	0.91274	0.88621	0.89343	0.54349

2) Deep learning model performance

Table 5 shows that, The CNN achieved the highest accuracy (87.13%) with balanced precision (84.49%), recall

Applied Information System and Management (AISM)
Volume 8, (2) 2025, p. 195–202
P-ISSN: 2621-2536; E-ISSN: 2621-2544; DOI: 10.15408/aism.v8i2.46658
©2025. The Author(s). This is an open acces article under cc-by-sa

(83.62%), and F1-score (84.05%), making it the most effective model. VGG16 also performed well with 85.79% accuracy. MLP and TabNet showed slightly lower accuracies (83.89% and 83.12%), with TabNet offering the best precision among tabular models. Overall, CNN provided the best classification performance.

Table 5.
Performance of DL Models

Model	Accuracy	Precision	Recall	F1-score
MLP	83.89%	79.05%	81.05%	80.03%
TabNet	83.12%	80.52%	78.06%	79.24%
CNN	**87.13%**	84.49%	83.62%	84.05%
VGG16	85.79%	83.20%	81.41%	82.30%

Given that the CNN model demonstrated superior performance, Fig. 6 illustrates the confusion matrix with the CNN model.

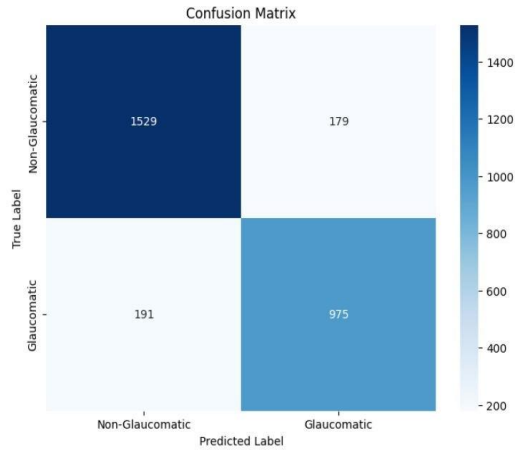

Fig. 6. Confusion matrix of CNN model.

C. Discussion

The superior performance of the YOLOv8 model in detecting the OD and OC can be attributed to its advanced architectural improvements over YOLOv5. YOLOv8's anchor-free design simplifies the detection process by removing the dependency on predefined anchor boxes, allowing the model to better detect objects with varying shapes and sizes—particularly useful for accurately localizing the OD and OC, which can present with subtle variations across images. Additionally, YOLOv8 leverages improved feature fusion through its decoupled head, enabling more precise localization and classification simultaneously. This enhancement contributes to the higher precision, recall, and mAP scores observed in our experiments, especially when compared to YOLOv5.

In the classification stage, the CNN model achieved the highest accuracy (87.13%), outperforming both traditional ML and other DL models. This can be explained by CNN's inherent ability to capture complex spatial patterns and relationships within the tabular feature space, particularly when these features are carefully structured to retain spatial dependencies like asymmetry, ISNT rule, and neuroretinal rim (NRR) thinning. CNN's convolutional layers excel at automatically learning hierarchical feature representations, which significantly boosts classification performance.

However, the CNN approach is not without limitations. Deep models like CNNs typically require substantial amounts of data to avoid overfitting, which can be a concern with smaller medical datasets. Additionally, CNNs demand more computational resources during training compared to simpler models like AdaBoost. On the other hand, AdaBoost demonstrated strong performance (83.22% accuracy) with much lower computational cost, making it suitable for deployment in resource-constrained settings.

The integration of ROI extraction also proved essential, as it allowed the models to focus on clinically relevant areas while reducing noise from irrelevant background regions. This preprocessing step directly improved classification outcomes by enhancing feature clarity.

D. Key Findings

YOLOv8 significantly outperformed YOLOv5 in OD and OC detection, thanks to its anchor-free architecture and superior feature fusion capabilities. CNN achieved the highest classification accuracy (87.13%), demonstrating its effectiveness in handling structured feature data extracted from fundus images. ROI extraction was a critical step, improving model efficiency and diagnostic accuracy, especially valuable in low-resource clinical environments. AdaBoost offered a computationally efficient alternative, achieving solid performance with lower resource demands. The proposed system successfully meets the research objective of providing an accurate, automated, and computationally practical solution for glaucoma screening by combining precise detection, clinically relevant feature extraction, and effective classification. This work underscores the impact of AI-driven pipelines in enhancing glaucoma diagnosis, offering a scalable, time-efficient tool that could be highly beneficial in clinical screening settings.

V. Conclusion

This study contributes to construct an automated framework for early glaucoma detection by integrating YOLOv8 for OD and OC localization with CNN for classification. Using a large-scale retinal fundus image dataset, the proposed system achieved a classification accuracy of 87.13%, demonstrating its effectiveness. A key contribution of this work is the incorporation of clinically relevant features, such as the ISNT rule and neuroretinal rim thinning, which enhance the diagnostic precision beyond standard image analysis. The framework offers a scalable, cost-effective solution suited for clinical environments, particularly benefiting low-resource settings where expert ophthalmologists may be scarce.

However, limitations include the dataset's limited demographic diversity and the reliance on bounding box annotations, which may not capture subtle optic nerve changes fully. Future research should focus on expanding dataset diversity, integrating multimodal clinical data, using polygon-based segment methods, and validating the system in real-world clinical settings. Developing a real-time deployment platform could further facilitate early glaucoma screening and help reduce the incidence of vision loss. Overall, this study contributes a significant step toward accessible, accurate automated glaucoma diagnosis.

REFERENCES

[1] A. Sallam *et al.*, "Early detection of glaucoma using transfer learning from pre-trained cnn models," in *2021 International Conference of Technology, Science and Administration (ICTSA)*, IEEE, 2021, pp. 1–5.

[2] M. Juneja *et al.*, "Automated detection of Glaucoma using deep learning convolution network (G-net)," *Multimed. Tools Appl.*, vol. 79, pp. 15531–15553, 2020.

[3] K. Allison, D. Patel, and O. Alabi, "Epidemiology of glaucoma: the past, present, and predictions for the future," *Cureus*, vol. 12, no. 11, 2020.

[4] Glaucoma Research Foundation, "Glaucoma Worldwide: A Growing Concern." 2024. [Online]. Available: https://glaucoma.org/articles/glaucoma-worldwide-a-growing-concern

[5] The Business Standard, "13 out of 100 people aged over 35 in Bangladesh suffer from glaucoma: Survey," *Bus. Stand.*, 2024, [Online]. Available: https://www.tbsnews.net/bangladesh/health/13-out-100-people-aged-over-35-bangladesh-suffer-glaucoma-survey-845786

[6] S. Ovreiu, I. Cristescu, F. Balta, and E. Ovreiu, "An exploratory study for glaucoma detection using densely connected neural networks," in *2020 International Conference on e-Health and Bioengineering (EHB)*, IEEE, 2020, pp. 1–4.

[7] M. J. Zedan, M. A. Zulkifley, A. A. Ibrahim, A. M. Moubark, N. A. M. Kamari, and S. R. Abdani, "Automated glaucoma screening and diagnosis based on retinal fundus images using deep learning approaches: A comprehensive review," *Diagnostics*, vol. 13, no. 13, Art. no. 2180, 2023, doi: 10.3390/diagnostics13132180.

[8] W. Liao, B. Zou, R. Zhao, Y. Chen, Z. He, and M. Zhou, "Clinical interpretable deep learning model for glaucoma diagnosis," *IEEE J. Biomed. Health Inform.*, vol. 24, no. 5, pp. 1405–1412, May 2020, doi: 10.1109/JBHI.2019.2949075.

[9] B. Turgut, "Pearls for Correct Assessment of Optic Disc at Glaucoma Diagnosis," *US Ophthalmic Rev.*, vol. 10, no. 2, pp. 104–110, 2017, doi: 10.17925/USOR.2017.10.02.104.

[10] A. Neto, J. Camara, and A. Cunha, "Evaluations of deep learning approaches for glaucoma screening using retinal images from mobile device," *Sensors*, vol. 22, no. 4, Art. no. 1449, 2022, doi: 10.3390/s22041449.

[11] T.-M. Kuang, C. J.-L. Liu, Y.-C. Ko, S.-M. Lee, C.-Y. Cheng, and P. Chou, "Distribution and associated factors of optic disc diameter and cup-to-disc ratio in an elderly Chinese population," *J. Chin. Med. Assoc.*, vol. 77, no. 4, pp. 203–208, 2014, doi: 10.1016/j.jcma.2014.01.006.

[12] A. Almazroa, S. Alodhayb, K. Raahemifar, and V. Lakshminarayanan, "An Automatic Image Processing System for Glaucoma Screening," *Int. J. Biomed. Imaging*, vol. 2017, pp. 1–9, 2017, doi: 10.1155/2017/4826385.

[13] R. Mahum, S. U. Rehman, O. D. Okon, A. Alabrah, T. Meraj, and H. T. Rauf, "A novel hybrid approach based on deep CNN to detect glaucoma using fundus imaging," *Electronics*, vol. 11, no. 1, Art. no. 26, 2021, doi: 10.3390/electronics11010026.

[14] S. Sreng, N. Maneerat, K. Hamamoto, and K. Y. Win, "Deep learning for optic disc segmentation and glaucoma diagnosis on retinal images," *Appl. Sci.*, vol. 10, no. 14, p. 4916, 2020.

[15] K. T. M. Han, P. Boonsieng, W. Kongprawechnon, P. Vejjanugraha, W. Ruengkitpinyo, and T. Kondo, "An Automated Framework for Screening of Glaucoma using Cup-to-Disc Ratio and ISNT Rule with a Support Vector Machine," *ResearchGate*, 2022, [Online]. Available: https://www.researchgate.net/publication/360299136_An_Automated_Framework_for_Screening_of_Glaucoma_using_Cup-to-Disc_Ratio_and_ISNT_Rule_with_a_Support_Vector_Machine

[16] A. A. Alqarni, S. H. Al-Harbi, and I. A. Subhan, "Creating an Early Diagnostic Method for Glaucoma Using Convolutional Neural Networks," *Preprint*, Mar. 2024, doi: 10.21203/rs.3.rs-4107145/v1.

[17] E. Maupin *et al.*, "Accuracy of the ISNT rule and its variants for differentiating glaucomatous from normal eyes in a population-based study," *Br. J. Ophthalmol.*, vol. 104, no. 10, pp. 1412–1417, 2020, doi: 10.1136/bjophthalmol-2019-315554.

[18] Sabari50312, "Fundus Pytorch Dataset." Kaggle, 2021. [Online]. Available: https://www.kaggle.com/datasets/sabari50312/fundus-pytorch

[19] L. K. Singh, Pooja, H. Garg, and M. Khanna, "Performance evaluation of various deep learning based models for effective glaucoma evaluation using optical coherence tomography images," *Multimed. Tools Appl.*, vol. 81, no. 19, pp. 27737–27781, 2022.

[20] L. J. Coan *et al.*, "Automatic detection of glaucoma via fundus imaging and artificial intelligence: A review," *Surv. Ophthalmol.*, vol. 68, no. 1, pp. 17–41, 2023.

[21] A. Almazroa, R. Burman, K. Raahemifar, V. Lakshminarayanan, and others, "Optic disc and optic cup segmentation methodologies for glaucoma image detection: A survey," *J. Ophthalmol.*, vol. 2015, 2015.

[22] A. Iwase, S. Sawaguchi, K. Tanaka, T. Tsutsumi, and M. Araie, "The Inferior, Superior, Temporal Rim Width Pattern (IST Rule) Detects Glaucoma in a Japanese Population," *J. Glaucoma*, vol. 31, no. 4, pp. 228–234, 2021, doi: 10.1097/IJG.0000000000001960.

[23] M. N. Bajwa *et al.*, "Two-stage framework for optic disc localization and glaucoma classification in retinal fundus images using deep learning," *BMC Med. Inform. Decis. Mak.*, vol. 19, pp. 1–16, 2019.

[24] E. Dervisevic, S. Pavljasevic, A. Dervisevic, and S. S. Kasumovic, "Challenges in early glaucoma detection," *Med. Arch.*, vol. 70, no. 3, p. 203, 2016.

[25] X. Huang *et al.*, "Detecting glaucoma from multi-modal data using probabilistic deep learning," *Front. Med.*, vol. 9, p. 923096, 2022.

[26] H. N. Veena, A. Muruganandham, and T. S. Kumaran, "A Review on the optic disc and optic cup segmentation and classification approaches over retinal fundus images for detection of glaucoma," *SN Applied Sciences*, vol. 2, no. 9, Aug. 2020, doi: 10.1007/s42452-020-03221-z.

[27] P. Das, S. R. Nirmala, and J. P. Medhi, "Detection of glaucoma using Neuroretinal Rim information," in *2016 International Conference on Accessibility to Digital World (ICADW)*, 2016, pp. 181–186. doi: 10.1109/ICADW.2016.7942538.

[28] X. R. Gao, F. Wu, P. T. Yuhas, R. K. Rasel, and M. Chiariglione, "Automated vertical cup-to-disc ratio determination from fundus images for glaucoma detection," *Sci. Rep.*, vol. 14, p. 4494, 2024, doi: 10.1038/s41598-024-55056-y.

[29] T.-M. Kuang, C. J.-L. Liu, Y.-C. Ko, S.-M. Lee, C.-Y. Cheng, and P. Chou, "Distribution and associated factors of optic disc diameter and cup-to-disc ratio in an elderly Chinese population☆," *Journal of the Chinese Medical Association*, vol. 77, no. 4, pp. 203–208, Mar. 2014, doi: 10.1016/j.jcma.2014.01.006.