African Vision and Eye Health
ISSN: (Online) 2410-1516, (Print) 2413-3183

Artificial intelligence-driven glaucoma screening in ophthalmology: The GlaucoNet deep learning framework

Check for updates

Authors:
Minakshi Kumar[1]
Deepak Dembla[1]
Vishal Goyal[2]

Affiliations:
[1]School of Computer Applications, Jaipur Engineering College and Research Centre (JECRC), Jaipur, India

[2]Department of Computer Applications, Punjabi University, Punjabi, India

Corresponding author:
Minakshi Kumar,
minakshi.agrawal@welingkar.org

Dates:
Received: 11 June 2025
Accepted: 29 Nov. 2025
Published: 25 Mar. 2026

How to cite this article:
Kumar M, Dembla D, Goyal V. Artificial intelligence-driven glaucoma screening in ophthalmology: The GlaucoNet deep learning framework. Afr Vision Eye Health. 2026;85(1), a1083. https://doi.org/10.4102/aveh.v85i1.1083

Copyright:
© 2026. The Author(s). Licensee: AOSIS. This work is licensed under the Creative Commons Attribution 4.0 International (CC BY 4.0) license (https://creativecommons.org/licenses/by/4.0/).

Read online:
Scan this QR code with your smart phone or mobile device to read online.

Background: Glaucoma is a significant cause of permanent blindness in the world, and it is usually unnoticed at its initial stages. Deep learning allows scaling flexible quality screening of retinal fundus images to detect and implement effective intervention at the appropriate time.

Aim: This study developed and evaluated GlaucoNet, an ensemble of MobileNetV2, InceptionV3, and ResNet50, for accurate glaucoma detection from retinal fundus images.

Setting: Publicly available data sets used include Online Retinal fundus Image database for Glaucoma Analysis (325 normal and 325 glaucoma images), and Automatic Classification of Retinal Images for Medical Assessment (309 normal and 396 glaucoma images).

Methods: To address the imbalance in classes, the process of data augmentation such as flipping, rotation, scaling and brightness was used. The data were separated into 5-fold cross-validation with patient-wise data separation training (70%), validation (10%) and testing (20%) to avoid data leakage. The early stopping that was done on validation loss, dropout (0.3) and L2 regularisation minimised overfitting in the branches of the convolutional neural network (CNN).

Results: The accuracy of GlaucoNet was 97.0, sensitivity 97.32, and specificity 96.64, with the following 95% confidence intervals of 96.19–97.8, 96.39–98.68 and 95.59–96.63 respectively. DeLong testing showed that the area under the curve of GlaucoNet (0.9698) was significantly greater compared to MobileNetV2, InceptionV3 and ResNet50 ($p < 0.05$) and proved the statistical dominance of the ensemble method.

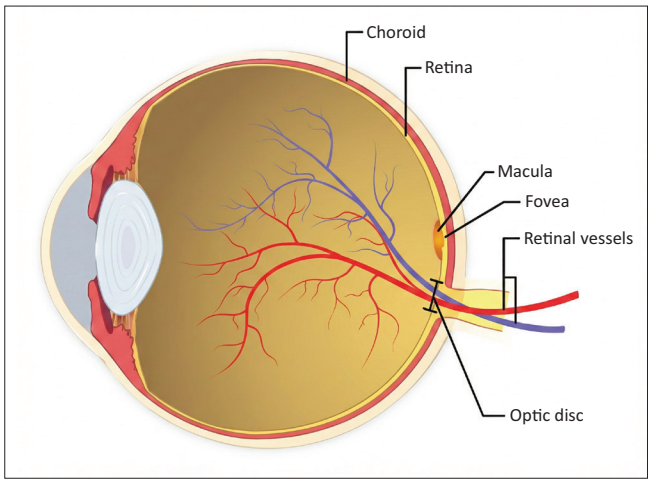
Conclusion: GlaucoNet achieved high accuracy in glaucoma detection, supporting its clinical screening potential.

Contribution: By leveraging the strengths of multiple CNN architectures, GlaucoNet offers a promising tool for enhancing early diagnosis and facilitating timely clinical intervention in ophthalmology.

Keywords: GlaucoNet; glaucoma; inception; ResNet; mobilenet; CNN; deep learning; transfer learning.

Introduction

Eye diseases such as glaucoma predominantly impact those aged between 40 and 80. Glaucoma is a leading cause of irreversible vision loss globally.[1] Numerous early-stage glaucoma patients are asymptomatic, resulting in many being undetected without routine clinical evaluation. If therapy is not administered, it may lead to enduring visual impairment and progressive optic nerve degeneration, so significantly diminishing quality of life.[2] Early anatomical abnormalities in the optic nerve head[3] can be found with fundus imaging, which records the posterior segment of the eye. Nevertheless, many underdeveloped areas have limited access to qualified ophthalmologists; thus, the progress of automated technology for early identification is necessary.[4] In this study, we have presented an ensemble deep learning method combining several models to improve glaucoma detection from fundus images by means of accuracy and dependability enhancement. Particularly helpful for medical image analysis, deep learning methods have shown great efficiency in identifying minute changes in retinal characteristics relevant to glaucoma diagnosis.[3,5,6] Developing effective glaucoma detection methods depends on a complete knowledge of the disorder. The main elements observed in a retinal fundus image are shown in Figure 1.[7] Maintaining ocular health depends on the complicated interaction between the anatomy



Source: Srinivasan PP, Kim LA, Mettu PS, Cousins SW. Fundus photography: Principles, applications and future directions. Retina Today. 2015;10(6):28–32

FIGURE 1: Anatomical features of the fundus.

and fluid dynamics of the human eye. Clear fluid in front of the eye; the aqueous humour is necessary to nourish ocular tissues and maintain intraocular pressure (IOP). Usually in balance, aqueous humour generation and elimination preserve normal IOP. When the glaucoma drainage system fails, fluid accumulates in the anterior chamber of the eye. Rising IOP can damage the optic nerve and impair vision and generate the characteristic glaucoma sign of optic cupping. Rising IOP in glaucoma can damage optic nerves and aggravate optic disc (OD) cupping.

Calculated by dividing the cup's diameter by the disc's total diameter, the cup-to-disc ratio (CDR) reveals the degree of glaucoma. A normal CDR value is around 0.5; raised levels indicate disease development.[8,9] An important worldwide health issue becoming more common is glaucoma. The open-angle glaucoma (OAG) and angle-closure glaucoma (ACG) are the two main forms; both have unique traits. Early glaucoma diagnosis is difficult because the illness usually shows no symptoms. This emphasises the need to put in place exact and efficient screening procedures.[10] Although earlier studies[11,12] show enormous potential for deep learning to detect glaucoma from fundus images, the field is still under development. We propose to build a GlaucoNet model combining the MobileNet, ResNet50, and Inception-v3 architectures in order to improve the accuracy and dependability of diagnosis. We want to provide an insightful analysis of several convolutional neural network (CNN) configurations for glaucoma classification, so it benefits the research community.

Review of literature

Recent developments in artificial intelligence have greatly increased the diagnosis of glaucoma. Moon et al.,[13] for instance, estimated visual field loss from swept-source optical coherence tomography (OCT) images using an Inception-ResNet-V2 model, thereby reporting improved performance in inferior retinal areas. Retinal pictures obtained on cell phones allowed Neto et al. to examine deep learning techniques. Their amazing classifying and segmenting accuracy came from using ResNet152 V2. Based on EfficientNet versions that routinely generate results with less parameters than standard models, Zarzà et al.[14] suggested a resource-efficient training technique based on Park et al.[15] who evaluated several designs and found, despite performance declining with illness severity, Inception-ResNet-v2 gave the most accurate estimations of visual field loss. Akbar et al.[16] identified early glaucoma from fundus images using DenseNet and DarkNet architectures, therefore obtaining strong accuracy and robustness. Gupta et al.[17] created a CNN-based glaucoma risk prediction system surpassing conventional clinical markers using confocal scanning laser ophthalmoscopy.

By combining different modalities, such as OCT and visual fields, Zhang et al.[18] underscored how artificial intelligence models based on ResNet and Visual Geometry Group (VGG) enhance clinician decision-making. Hu et al.[19] creatively used fine-tuned Bidirectional Encoder Representations from Transformers (BERT) models to anticipate surgical demand relying on unstructured clinical notes, hence leveraging natural language processing.

Mobile health potential was underscored by high accuracy and user satisfaction of a smartphone-based diagnostic app created by Guo et al.[20] Hemelings et al.[21] produced a generalisable fundus-based glaucoma classifier by using transfer and active learning. Long et al.[22] cautioned against complexity and overfitting and suggested simplified models, including penalised support vector machine (SVM) for clinical translation. Finally, Kumari et al.[23] presented an ensemble model with the possible ability to integrate the capabilities of numerous CNNs. Kumari et al.[23] generated fusion models that yielded outstanding prediction results (F1: 0.745, area under the curve [AUC]: 0.899) by combining textual and structural data. Sharifi et al.[24] applied supervised machine learning using Random Forest and Decision Trees on visual field, CDR, and blood pressure data. Diaz-Alemán et al.[25] evaluated image-based deep learning on ganglion cell layer (GCL) pictures using ResNet101 and Shufflelet and attained exact early detection. In using insurance claims, Bremond-Gignac et al.[26] projected open-angle glaucoma with a 79% accuracy rate. Li et al.[5] predicted glaucoma progression in retinal pictures using deep learning; their Attention-based Glaucoma Convolutional Neural Network (AG-CNN) model[27] enhanced detection by using attention-based subnets. Hemelings et al.[21] developed a regression model with AUCs spanning 0.854–0.998 using 11 datasets. Gupta et al.[28] reported a class-balanced network using latent structure discovery to identify early visual symptoms. Hu et al.[29] created GLaucoma forecast transformer based on Irregularly saMpled fundus images (GLIM)-Net, which performed remarkably well to predict chronic glaucoma using consecutive fundus pictures. Yoon et al.[30] examined relationships between glaucoma and the oral flora using machine learning. Juneja et al.[31] improved optic disc (OD) and cup segmentation using CNNs, therefore raising the diagnosis's accuracy. Kamal et al.[32] used explainable artificial

intelligence (XAI) with models including Adaptive Neuro-Fuzzy Inference System (ANFIS) and Superpixel-based Local Interpretable Model-agnostic Explanations (SP-LIME) to improve risk assessments. Akter et al.[33] proposed a cross-sectional feature from Optical Coherence Tomography (OCT) scans, therefore raising diagnosis accuracy to 97%. Lee et al.[34] used machine learning to forecast long-term glaucoma depending on optic nerve characteristics and early ocular data. Nakahara et al.[35] highly accurately identified glaucoma using fundus images from smartphones. Gao et al.[36] improved primary open-angle glaucoma (POAG) prediction utilising genetic data by means of polygenic risk scores (PRS). Devecioglu et al.[37] first presented self-optic nerve networks (self-ONNs) for real-time glaucoma detection in digital fundus images. Jha et al.[38] and Dubey et al.[39] separately examined ensemble and transfer learning, respectively, in plant disease identification using crossover applications to medical diagnostics. A study[40] claims that machine learning could assist in creating glaucoma models and enhance the diagnosis of cardiovascular disease (CVD). In recent studies, DeepGlauNet[41] combined CNN feature maps with attention mechanisms for improved OD–cup differentiation, while G1020[42] achieved a high AUC on the Retinal Image Database for Optic Nerve Evaluation (RIM-ONE) dataset through fine-tuned transfer learning.

Research methods and design

Dataset overview: Study on glaucoma detection utilising public datasets

Two fundamental glaucoma datasets – ORIGA and ACRIMA – are used in this work. The ORIGA dataset used in the current research has 325 normal and 325 glaucoma images, and the ACRIMA dataset has 309 normal and 396 glaucoma images. Developed by the Singapore Eye Research Institute, ORIGA comprises 650 annotated fundus photos (2004–2007) labelled for glaucoma or healthy condition with OD, optic cup (OC) outlines, and CDR values.[43] Originating from a class imbalance favouring healthy samples, ORIGA is still important for segmentation and classification. Because of consistent disc-centred alignment, ACRIMA consists of 705 centred OD images (309 normal, 396 glaucoma), providing simplicity of preprocessing. It does not, however, have OD or OC segmentation markings.[44] Images from both sets were downsized to 224 × 224 pixels, and the quality was improved with Python APIs. Images preprocessed to align and normalise CDR comprise 650 ORIGA and 705 ACRIMA samples in total.

Glaucoma image data augmentation

An effective solution to the lack of medical photos is data augmentation. To address the imbalance, data augmentation techniques such as flipping, rotation, scaling, and brightness variation were applied to expand the minority class samples, ensuring more uniform representation during training. This approach boosts the model's performance by making it better at handling new data and helps to prevent it from becoming too reliant on just a few images.

With Keras's Image Data Generator specifically, we accomplished the following:

- Random rotations, horizontal flips, and zooms allow geometric changes in scale and direction.
- Changing brightness, contrast, and saturation will help you to replicate various lighting conditions.

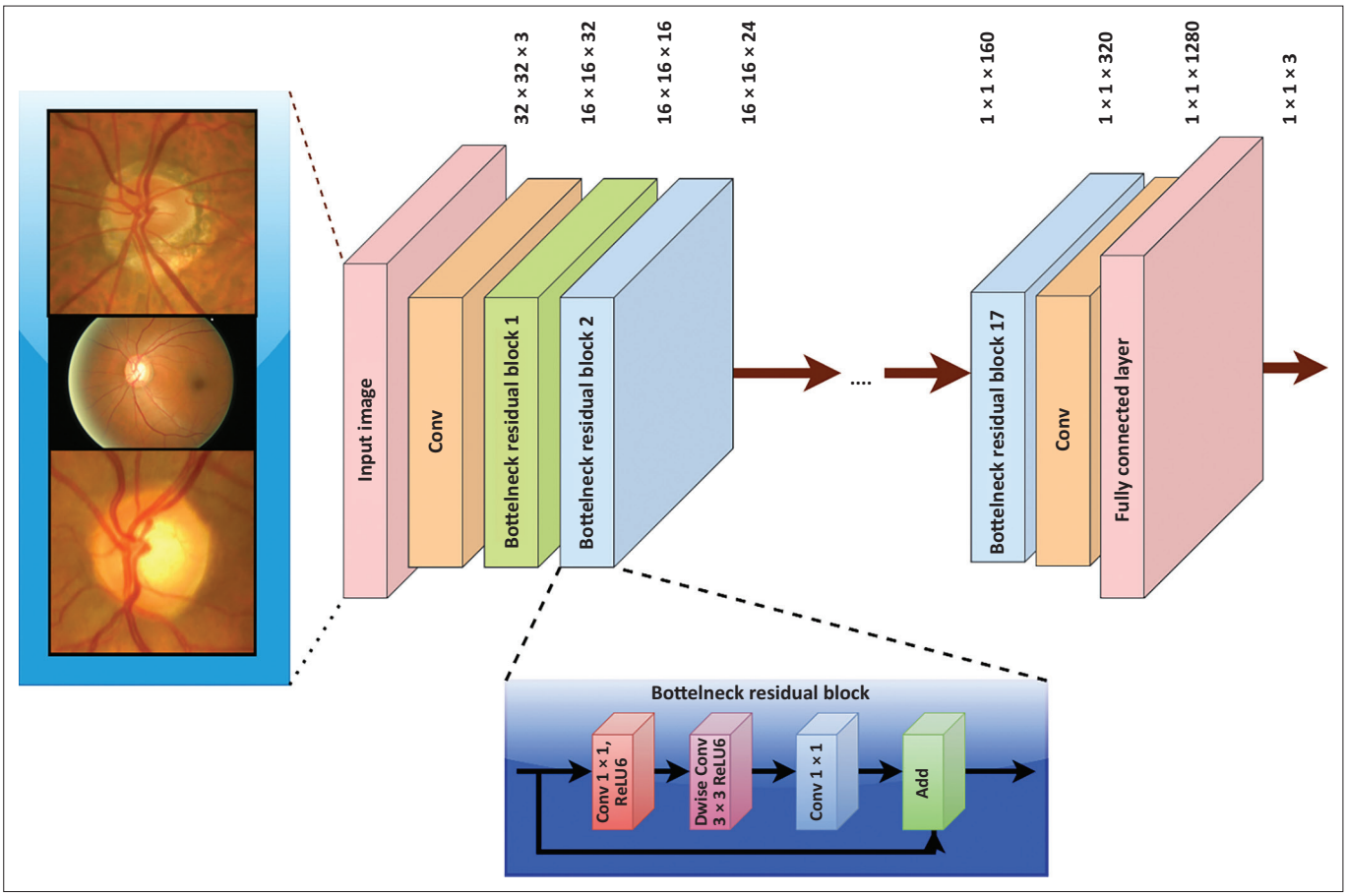
These techniques sought to enable the model to acquire discriminative characteristics from fundus images and generalise successfully to new situations.[45] The current research uses *k*-fold cross-validation, with a value of $k = 5$; the combined dataset was divided into 70% training, 10% validation, and 20% testing. Images from the same patient were maintained within the same subset to prevent data leakage and maintain subject-level independence. In addition to augmentation, early stopping (based on validation loss), dropout layers (rate = 0.3), and L2 regularisation were incorporated into each CNN branch to mitigate overfitting and enhance model generalisation.

MobileNetV2 architecture for glaucoma disease prediction

Particularly fit for the image classification process, including glaucoma diagnosis, MobileNetV2 is a quite powerful CNN design. Figure 2 shows the glaucoma disease prediction architecture of MobileNetV2. This model is ideal for environments and uses limited resources because the depth-wise separable convolutions of it lower processing cost and model size while preserving accuracy. The architectural inverted residual blocks and linear bottlenecks help to extract features, therefore facilitating the identification of tiny glaucoma markers in retinal pictures. MobileNetV2 enables clinicians and researchers to locate exact and instantaneous glaucoma.

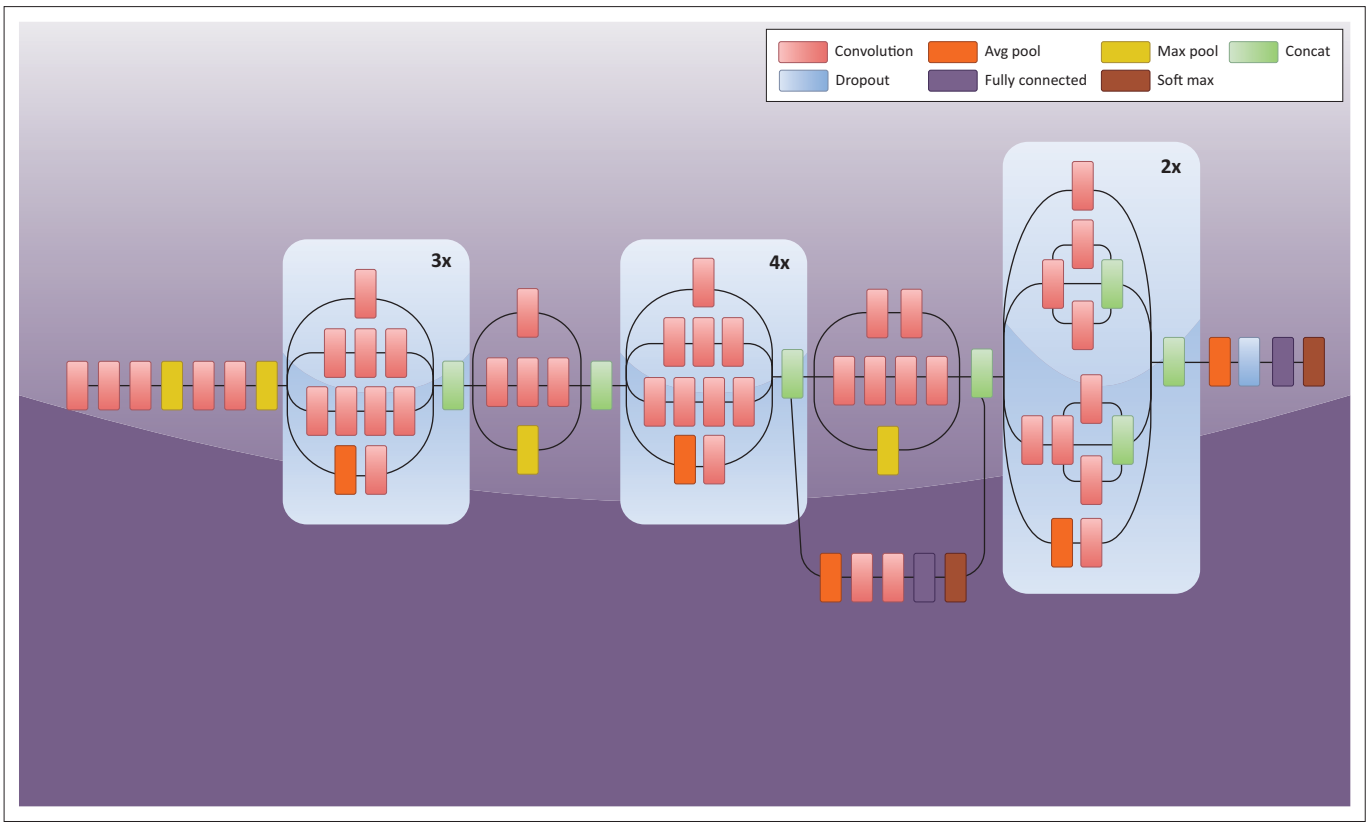
Inception V3 architecture for glaucoma disease prediction

Built mostly for image classification, with an eye on retinal scan glaucoma diagnosis, an advanced CNN structure known as Inception V3 was built. The design is imaginative and makes use of numerous quite crucial elements. Inception V3 architecture for glaucoma disease (Figure 3) predicts the capacity of inception modules to simultaneously process images at multiple sizes; it lets one gather a spectrum of glaucoma-related properties. Through division of large convolutions into smaller ones, factorised convolutions reduce computational complexity without sacrificing performance. Auxiliary classifiers and intermediary layers improve training by adding extra gradient signals, hence enhancing the resilience of the model. Strided convolutions and max pooling help one to shrink the grid. These techniques significantly reduce feature map sizes without sacrificing vital information. These architectural details taken together increase the accuracy of Inception V3 in retinal image feature extraction and glaucoma case classification.



Conv, donvolutional layer; Dwise, depthwise separable convolutions; ReLU6, rectified linear unit.

FIGURE 2: MobileNetV2 architecture for glaucoma disease prediction.



Avg, average; Max, maximum; Concat, concatenation.

FIGURE 3: Inception V3 architecture for glaucoma disease prediction.

ResNet V2 architecture for glaucoma disease prediction

ResNet V2, a modified version of Residual Network architecture, is well recognised for its remarkable depth and high performance in applications related to image classification. Architecture's pioneering design effectively tackles the issue of the vanishing gradient problem, hence facilitating the training of very deep neural networks.

ResNet V2 is well-suited for medical image analysis, namely, in predicting glaucoma disease. It excels at properly detecting tiny characteristics in retinal images. Figure 4 shows ResNet V2 architecture for glaucoma disease prediction.

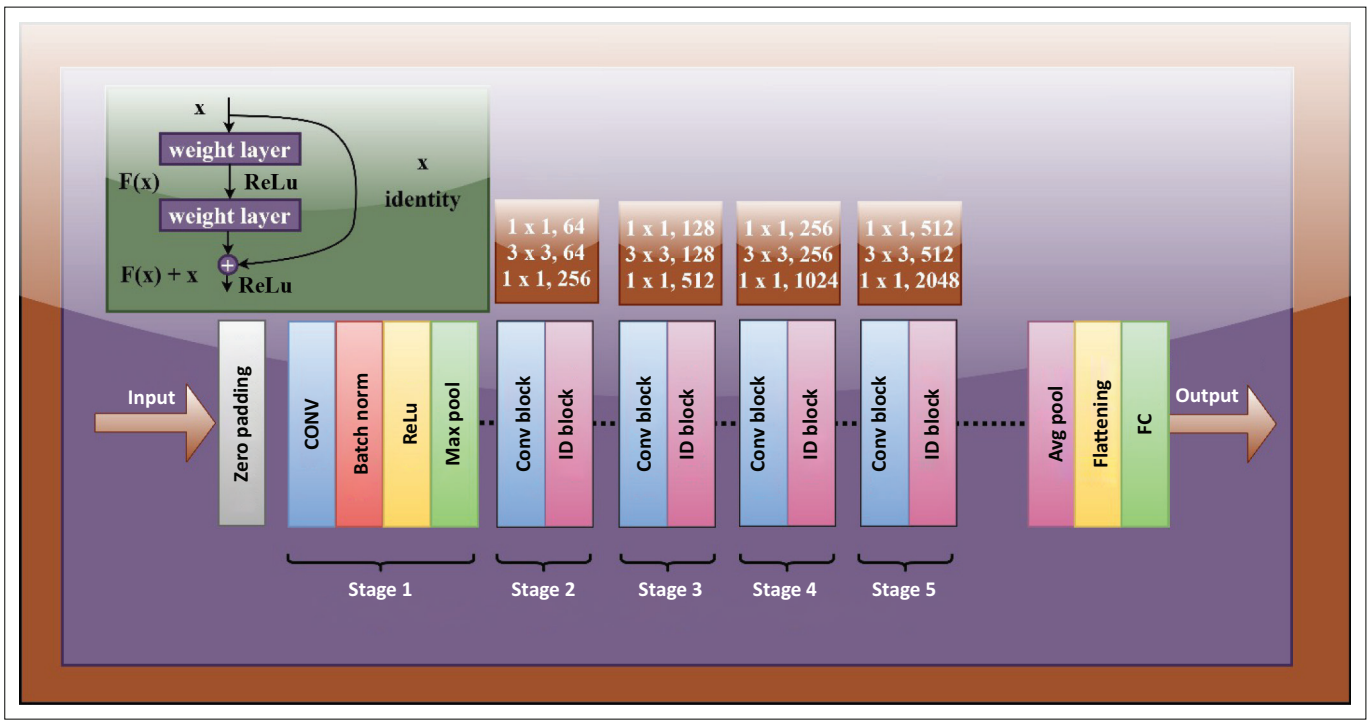
Architectural components

The architectural components are:

- Residual Blocks: ResNet V2's fundamental building block, Residual Blocks, uses identity mapping to address disappearing gradients. Every block comprises ReLU activation functions and batch normalisation (BN) accompanying a series of convolutional layers. An equivalent shortcut connection then follows to add the block's input to its output. Including this skip link helps the gradients to flow across the network, hence supporting the training of more complicated designs.
- ResNet V2 enhances the original ResNet by putting BN and ReLU activation before rather than after the weight layers in the residual blocks. Especially helpful when working with complex patterns in retinal images, the pre-activation method improves the stability and convergence of training.
- ResNet V2 architecture sometimes uses a bottleneck design in its residual blocks, with three layers: a 1 × 1 convolution for dimensionality reduction, a 3 × 3 convolution, and yet another 1 × 1 convolution for dimensionality restoration. This configuration reduces the processing load and lets one employ deeper networks without running too much computational cost.
- ResNet V2 substitutes the worldwide average pooling before the last softmax layer for the traditional completely linked layers. These steps are taken by lowering the number of parameters, therefore preventing overfitting, and making sure the network focuses on the most important traits for classification.

Proposed GlaucoNet Architecture for Glaucoma Disease Prediction

Combining three strong deep learning models – MobileNetV2, Inception, and ResNet – the method presents a whole solution for glaucoma identification. The process starts with data collected from the A CRIteria-based Retinal Image Database for Glaucoma Analysis (ACRIMA) and Online Retinal Fundus Image Database for Glaucoma Analysis (ORIGA) databases. After that, the data are compiled and divided into several sets for needs in training, validation, and testing. Among data augmentation methods, rotating, shifting, zooming, sharing, and channel shifting improve the quality of the training set. Each model is then trained with the better data, and thereafter maintained. The GlaucoNet model is formed by the average of the three models' forecasts. One can assess the model by means of the computation of evaluation criteria, including accuracy, sensitivity, specificity, and AUC. The GlaucoNet model is



CONV, convolution (layer); Conv, convolution (layer); ReLu, rectified linear unit; Max, maximum; Avg, average; FC, fully connected (layer).

FIGURE 4: ResNet V2 architecture for glaucoma disease prediction.

finally used to identify glaucoma by means of real-time predictions on newly acquired fundus images, thereby offering a rather accurate and efficient approach. Figure 5 shows the working idea of the proposed GlaucoNet model.

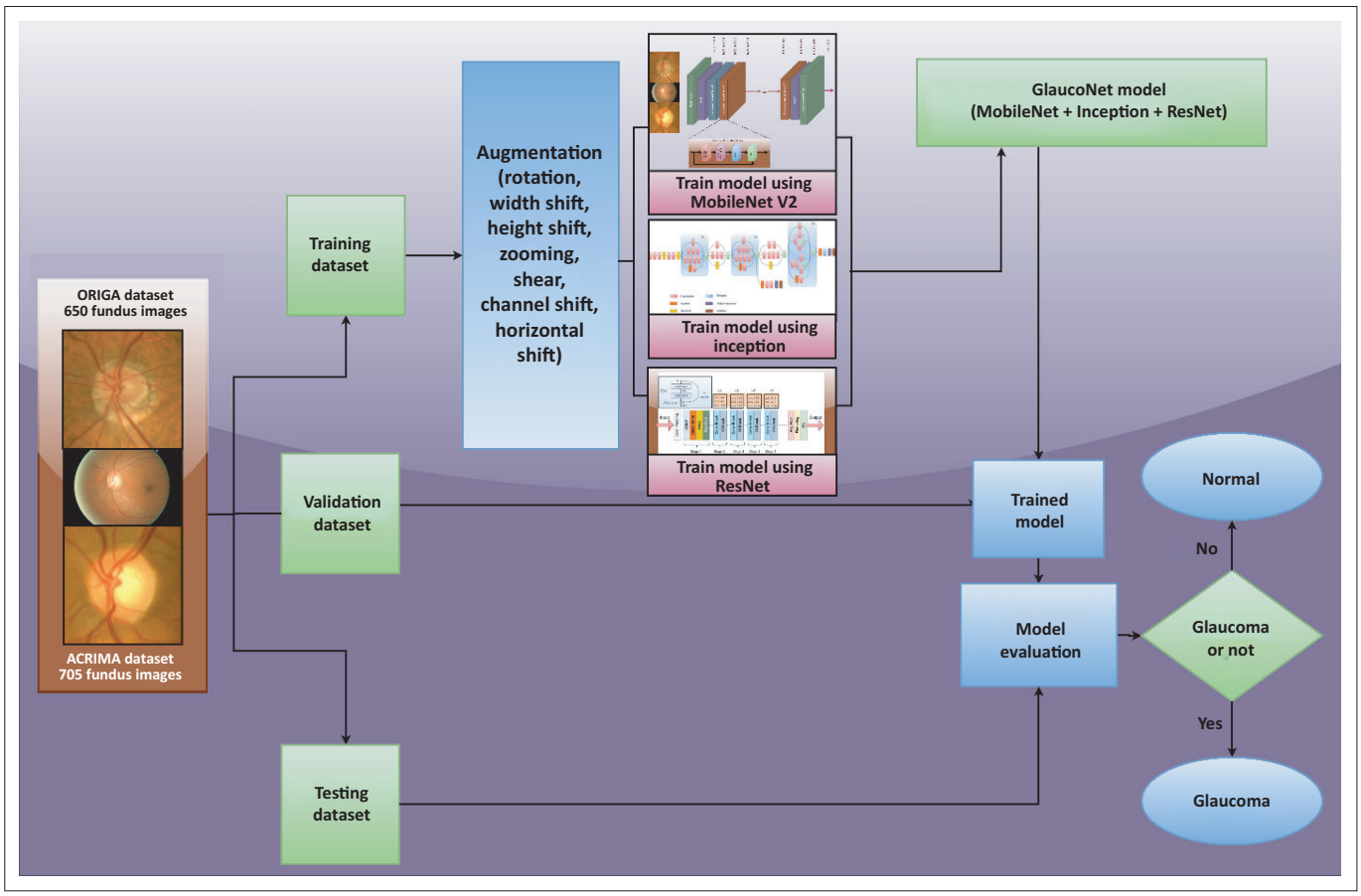
Predictions from MobileNetV2, InceptionV3, and ResNet50 are combined using a weighted averaging strategy, where each model's contribution is proportional to its validation accuracy. This weighted fusion ensures a balanced ensemble performance.

Statistical validation

To assess the robustness of model performance, 95% confidence intervals (CIs) were calculated for accuracy, sensitivity, specificity, and AUC values using bootstrap resampling (n = 1000). In addition, pairwise AUC comparisons between the proposed GlaucoNet model and the individual CNNs (MobileNetV2, InceptionV3, and ResNet50) were conducted using the DeLong test, providing statistical evidence of significance. A p-value of less than 0.05 was considered statistically significant.

Algorithm for the proposed GlaucoNet model

Begin
Step 1: Data Collection and Preparation
1.1 ORIGA_dataset ← LOAD("ORIGA")
1.2 ACRIMA_dataset ← LOAD("ACRIMA")
1.3 full_dataset ← COMBINE(ORIGA_dataset, ACRIMA_dataset)
1.4 (training_dataset, validation_dataset, testing_dataset) ← SPLIT(full_dataset, ratios = [0.7, 0.1, 0.2])
Step 2: Data Augmentation
2.1 ∀ x ∈training_dataset do
2.2 AUGMENTED_training_dataset ← {x_rotated, x_width_shifted, x_height_shifted, x_zoomed, x_sheared, x_channel_shifted, x_horizontal_shifted}
Step 3: Model Training
3.1 Train MobileNet V2 Model
MobileNetV2_model ← INITIALISE_MobileNetV2() TRAIN(MobileNetV2_model, AUGMENTED_training_dataset) SAVE(MobileNetV2_model, "MobileNetV2_model.h5")
3.2 Train Inception Model
Inception_model ← INITIALISE_Inception() TRAIN (Inception_model, AUGMENTED_training_dataset) SAVE(Inception_model, "Inception_model.h5")
3.3 Train ResNet Model
ResNet_model ← INITIALISE_ResNet() TRAIN(ResNet_model, AUGMENTED_training_dataset) SAVE(ResNet_model, "ResNet_model.h5")
Step 4: GlaucoNet Model Formation
4.1 MobileNetV2_model ← LOAD("MobileNetV2_model.h5")



ORIGA, Online Retinal Fundus Image Database for Glaucoma Analysis; ACRIMA, A CRIteria-based Retinal Image Database for Glaucoma Analysis.

FIGURE 5: Proposed GlaucoNet model for glaucoma disease prediction.

4.2 Inception_model ← LOAD("Inception_model.h5")
4.3 ResNet_model ← LOAD("ResNet_model.h5")
4.4 GlaucoNet_model(x) ← (PREDICT(MobileNetV2_model, x) + PREDICT(Inception_model, x) + PREDICT(ResNet_model, x)) / 3
Step 5: Model Evaluation
5.1 true_positive, true_negative, false_positive, false_negative ← 0, 0, 0, 0
5.2 ∀x ∈testing_dataset do y_pred ← GlaucoNet_model(x)
y_true ← TRUE_LABEL(x)
IF y_pred = y_true = 'Glaucoma' THEN true_positive ← true_positive + 1
ELSE IF y_pred = y_true = 'Normal' THEN true_negative ← true_negative + 1
ELSE IF y_pred = 'Glaucoma' AND y_true = 'Normal' THEN false_positive ← false_positive + 1
ELSE IF y_pred = 'Normal' AND y_true = 'Glaucoma' THEN false_negative ← false_negative + 1
5.3 ACC ← (true_positive + true_negative) / (true_positive + true_negative + false_positive + false_negative)
5.4 SN ← true_positive / (true_positive + false_negative)
5.5 SP ← true_negative / (true_negative + false_positive)
5.6 AUC ← CALCULATE_AUC(ROC_curve)
5.7 PRINT "Accuracy: ", ACC × 100, "%"
5.8 PRINT "Sensitivity: ", SN × 100, "%"
5.9 PRINT "Specificity: ", SP × 100, "%"
5.10 PRINT "AUC: ", AUC
Step 6: Deployment and Prediction
6.1 WHILE new_fundus_image IS AVAILABLE DO y_pred ← GlaucoNet_model(new_fundus_image)
IF y_pred = 'Glaucoma' THEN PRINT "Glaucoma detected" ELSE
PRINT "Normal"
End

Performance evaluation metrics

Accuracy, sensitivity, and specificity are assessed in the GlaucoNet model. Understanding these measurements calls for the following classifications: A True Positive (TP) occurs when the model accurately spots glaucoma in a picture. A True Negative (TN) is when the model correctly detects a healthy image. A False Positive (FP) is misclassifying a healthy picture as glaucoma-induced. When the model misclassifies glaucoma-impacted pictures as healthy, a False Negative (FN) results.

The accuracy ratio counts the frequency of accurate estimations against overall projections.

Sensitivity evaluates the glaucoma-detecting accuracy of the model (Equation 1):

$$\text{Sensitivity} = TP / (TP + FN). \quad [\text{Eqn 1}]$$

Specificity gauges if the model can identify healthy examples (Equation 2):

$$\text{Specificity} = TN / (TN + FP). \quad [\text{Eqn 2}]$$

The F1-score strikes a compromise between recall and accuracy (Equation 3):

$$\text{F1-score} = 2 * TP / (2 * TP + FP + FN) \quad [\text{Eqn 3}]$$

For model evaluation, the dataset was split into training (70%), validation (10%), and testing (20%) subsets. Every image was shrunk and centred on the OD to preserve uniformity.

Ethical considerations

This study used publicly available, de-identified retinal image datasets. Ethical approval was not required as no identifiable human data were involved. The research adhered to relevant data protection and ethical guidelines.

Results

Analysis of the precision-recall curve for glaucoma disease prediction

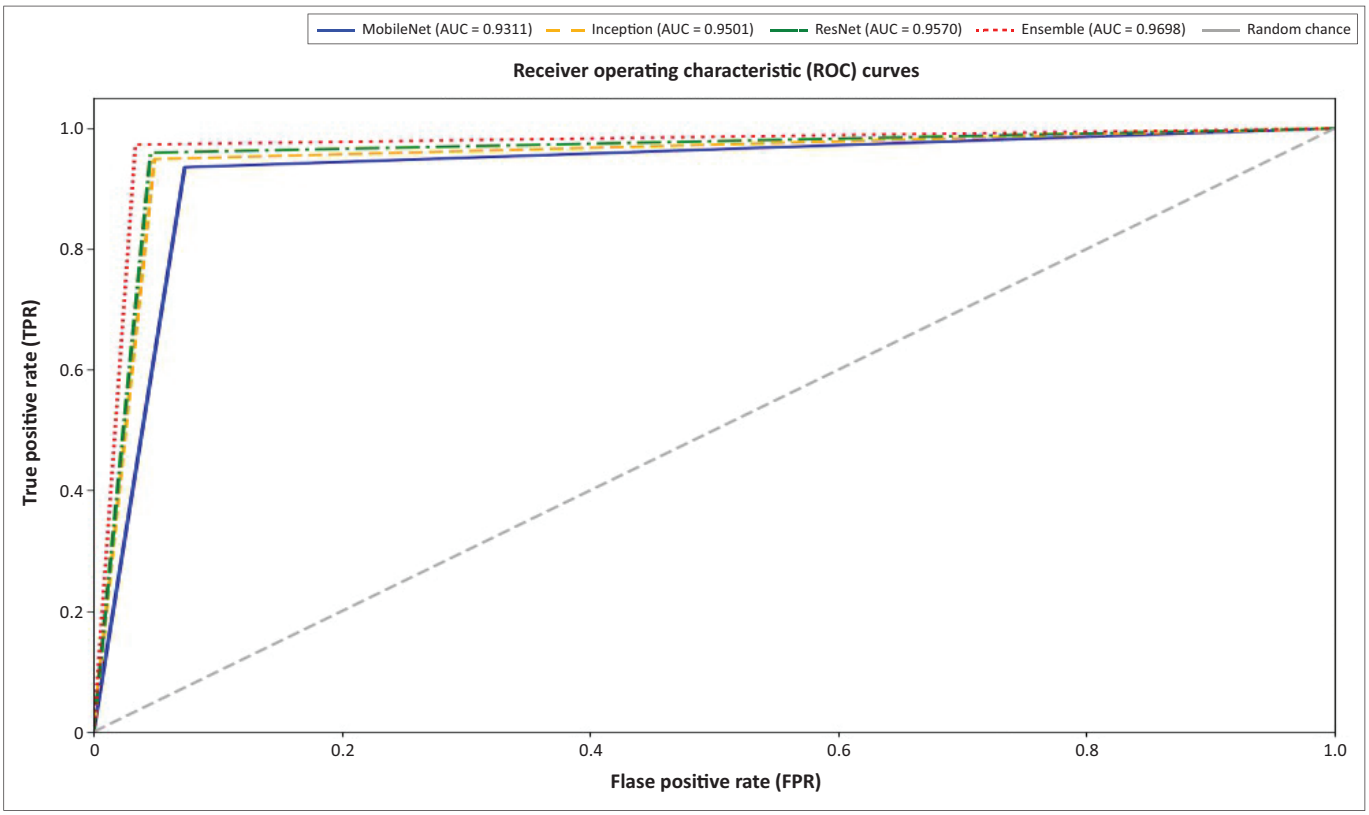
The Precision-Recall (PR) curve demonstrates that the proposed GlaucoNet model outperforms other models in predicting glaucoma disease, with an AUC of 0.9790 as shown in Figure 6. This indicates that the GlaucoNet model successfully achieves precision and recall. The ResNet model follows with an AUC of 0.9705, then Inception with an AUC of 0.9665, and MobileNet with an AUC of 0.9528. Maximising precision reduces the occurrence of FP, which in turn decreases redundant patient concern and testing.

On the other hand, maximising recall guarantees that the majority of glaucoma cases are identified, hence limiting the chances of untreated disease progression. The AUC values, ranging from 0.9528 to 0.9790, demonstrate that all models perform much better than random chance. Among them, the proposed GlaucoNet method is the most appropriate for accurate glaucoma prediction, which has the potential to enhance patient outcomes. The GlaucoNet model achieved an overall accuracy of 97.0% (95% CI: 96.1% – 97.8%), sensitivity of 97.32% (95% CI: 96.3% – 98.0%), and specificity of 96.64% (95% CI: 95.5% – 97.5%).

Using the DeLong test, the AUC of GlaucoNet (0.9698) was found to be significantly higher than that of MobileNetV2 (p = 0.021), InceptionV3 (p = 0.034), and ResNet50 (p = 0.041), confirming the statistical superiority of the ensemble approach.

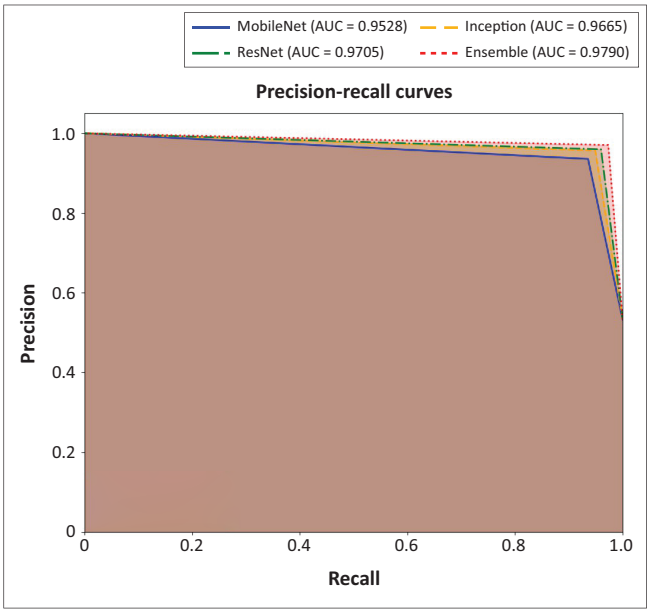
Analysis of receiver operating characteristic curve for glaucoma disease prediction

Figure 7 shows the receiver operating characteristic (ROC) curves for models of glaucoma prediction by stressing the trade-off between false positive rate (FPR) and true positive rate (TPR). These curves define most of the analysis of diagnostic performance. With the best AUC (0.9698), the



AUC, area under the curve.

FIGURE 6: Precision-recall curve for glaucoma disease prediction.



AUC, area under the curve.

FIGURE 7: Receiver operating characteristic curve for glaucoma disease prediction.

proposed GlaucoNet model demonstrated better general accuracy. ResNet (0.9570), Inception (0.9501), and MobileNet (0.9311) also did well, even if MobileNet showed much less efficiency.

Every model far exceeded the random estimate with AUCs much over 0.5. GlaucoNet is the most efficient for glaucoma detection because it shows the highest mix of sensitivity and specificity. Early diagnosis and treatment depend on maximising TPR and lowering FPR, therefore helping to stop vision loss and illness development.

Analysis of the confusion matrix for glaucoma disease prediction

The confusion matrices for the glaucoma disease prediction models, as shown in Figure 8, show that the proposed GlaucoNet model achieves the highest accuracy (0.9698), precision (0.9706), and recall (0.9732), followed by ResNet with an accuracy of 0.9570, precision of 0.9598, and recall of 0.9598, Inception with an accuracy of 0.9501, precision of 0.9567, and recall of 0.9492, and MobileNet with an accuracy of 0.9311, precision of 0.9356, and recall of 0.9356. These metrics highlight the proposed GlaucoNet model's superior ability to balance TP and TN predictions, making it the most effective for early glaucoma detection and prevention.

To provide deeper insight into model decision behaviour, a confusion matrix was generated for all CNN models and the ensemble GlaucoNet framework. The confusion matrix (Figure 8) illustrates the distribution of TP, TN, FP, and FN. GlaucoNet achieved the most balanced performance, showing the highest TP and TN rates with minimal FP and FN errors compared to MobileNetV2, InceptionV3, and ResNet50. Further analysis of misclassified cases revealed that FPs were primarily because of OD irregularities or illumination artefacts resembling glaucomatous cupping, while FNs were associated with early-stage glaucoma images exhibiting

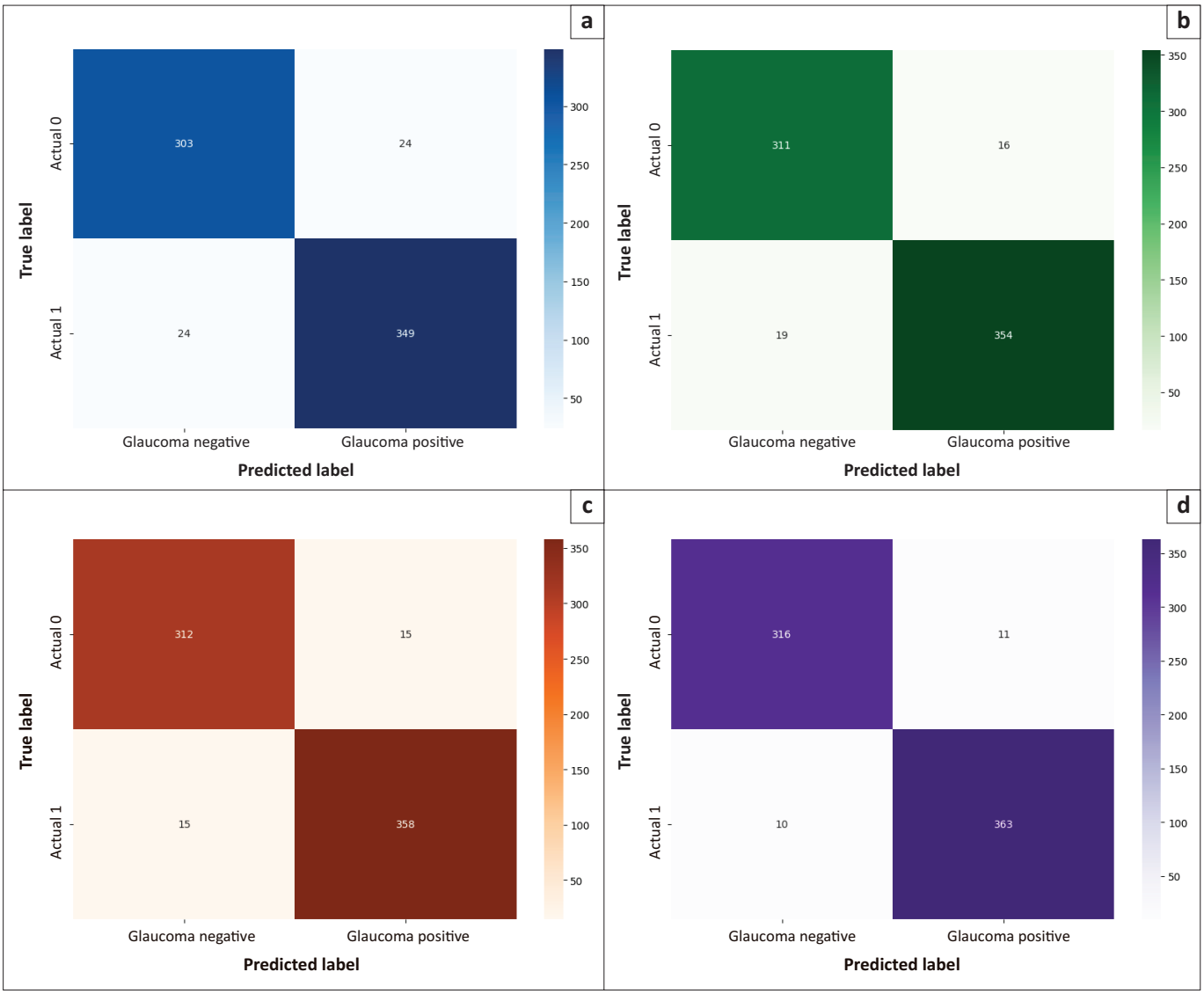


FIGURE 8: Confusion matrix for glaucoma disease prediction: (a) MobileNet, (b) Inception, (c) ResNet, and (d) Ensemble.

TABLE 1: Performance comparison of baseline convolutional neural networks and proposed GlaucoNet ensemble model on ORIGA and ACRIMA datasets.

Model	%				AUC
	Accuracy	Sensitivity	Specificity	Precision	
MobileNetV2	93.14	93.56	92.66	93.56	0.9311
InceptionV3	95.00	94.91	95.11	95.68	0.9529
ResNet50	95.71	95.98	95.42	95.98	0.9598
GlaucoNet (Proposed)	**97.00**	**97.32**	**96.64**	**97.06**	**0.9698**

ORIGA, Online Retinal Fundus Image Database for Glaucoma Analysis; ACRIMA, A CRIteria-based Retinal Image Database for Glaucoma Analysis; AUC, area under the curve.

subtle morphological changes. This analysis highlights the interpretability and practical limitations of AI-based glaucoma screening systems in real-world clinical contexts.

As shown in Table 1, the proposed GlaucoNet model achieved the highest accuracy (97.0%), sensitivity (97.32%), and specificity (96.64%), exceeding the individual performances of MobileNetV2, InceptionV3, and ResNet50. This consistent improvement across all evaluation metrics demonstrates the ensemble's superior ability to balance diagnostic precision and recall.

Comparison of the proposed GlaucoNet model with other existing models

The Proposed GlaucoNet Ensemble model outperforms previous methods on many datasets (Table 2). GlaucoNet's accuracy of 0.97 surpasses U-Net (0.88)[46] and VGG19 (0.8365).[47] The system's specificity (SP) and sensitivity (SN) are also high, at 0.9664 and 0.9732, indicating a better detection capacity. With an AUC of 0.9719, GlaucoNet outperforms GoogleNet (0.9277)[47] and DENet (0.9083).[47] In addition, GlaucoNet's F1 Score (0.9706) shows its ability to balance precision and recall, which is crucial in medical diagnosis. GlaucoNet Ensemble makes considerable gains in accuracy, robustness, and diagnostic reliability, surpassing current state-of-the-art models.

Limitations

Despite the fact that GlaucoNet showed better results in detecting glaucoma at its early stages, some shortcomings should be considered. The complexity of integrating the various outputs of different models can result in risks of

TABLE 2: Comparison of the proposed GlaucoNet model with other existing models.

Reference number	Methodology	Accuracy	SP	SN	AUC	F1 score	Precision	Dataset
Gómez-Valverde, Antón, Fatti et al.[47]	U-Net	0.88	0.86	0.91	0.93	-	-	RIM-ONE, DRISHTI
Serener, Serte[48]	VGG19	0.84	0.89	0.84	0.92	-	-	RIM-ONE
	ResNet	0.85	0.86	0.84	0.93	-	-	DRISHTI
	GoogleNet	0.86	0.83	0.89	0.93	-	-	ESPERANZA
	DENet	0.85	0.88	0.81	0.91	-	-	-
Jiang, Duan, Cheng et al.[49]	ResNet	0.86	0.93	0.21	-	-	-	RIM-ONE
	GoogleNet	0.85	0.91	0.29	-	-	-	-
Jha, Dembla, Dubey[39]	Inter-GD	-	-	-	-	79.90	81.40	I-OAD-B, DRISHTI
		-	-	-	-	77.80	79.90	I-OAD-B, ORIGA
Gampala, Maram, Vigneshwari, Cristin[40]	Joint-RCNN	-	-	-	0.90	-	-	ORIGA, SCES
Lavric, Petrariu, Havriliuc, Coca E[41]	Deep Neurofuzzy Network	92.21	92.34	93.42	-	-	-	Private
	VGG16 + LSTM	94.00	89.90	86.00	-	-	-	-
	ResNet50 + LSTM	88.00	83.00	79.00	-	-	-	-
Proposed Glauco Net-CNN model	MobileNet	0.93	0.93	0.94	0.94	0.94	-	ORIGA and ACRIMA
	Inception	0.95	0.95	0.95	0.95	0.96	-	-
	ResNet	0.96	0.95	0.96	0.96	0.96	-	-
	GlaucoNet	0.97	0.97	0.97	0.97	0.97	-	-

Note: Please see the full reference list of the article, Kumar M, Dembla D, Goyal V. Artificial intelligence-driven glaucoma screening in ophthalmology: The GlaucoNet deep learning framework. Afr Vision Eye Health. 2026;85(1), a1083. https://doi.org/10.4102/aveh.v85i1.1083, for more information.

CNN, convolutional neural network; SP, specificity; SN, sensitivity; AUC, area under the curve; VGG, Visual Geometry Group; GD, glaucoma detection; RCNN, region-based convolutional neural network; LSTM, long short-term memory; ORIGA, Online Retinal Fundus Image Database for Glaucoma Analysis; ACRIMA, A CRIteria-based Retinal Image Database for Glaucoma Analysis.

overfitting or biasness with ensemble models such as GlaucoNet. Although the ensemble is more accurate and robust than the individual models, there is a trade-off between the part-based (part CNNs) and the whole-based (ensemble) reasoning, and this trade-off has philosophical and practical difficulties. In particular, the predictive choices that the ensemble makes can be less interpretable, so it may be difficult to tell what model or features have the strongest impact on the final classification. Also, the research employed publicly available datasets of a limited sample size, something that could limit generalisability. More extensive and heterogeneous datasets that would be further validated (with further exploration of interpretability approaches) would enhance the applicability of the models to clinical use.

Conclusion and future scope

The present work shows fairly superior performance of the proposed GlaucoNet model for early glaucoma detection using fundus images. GlaucoNet combines three CNN architectures: MobileNetV2, InceptionV3, and ResNet50 to surpass any one model. Its accuracy of 97.0% exceeded MobileNetV2 (93.14%), InceptionV3 (95.0%), and ResNet50 (95.71%). GlaucoNet generated remarkable results with a sensitivity rate of 97.32% and a specificity of 96.64%. It is thus excellent in showing both accurate positives and negatives. Its Precision-Recall AUC score was 0.9790, better than ResNet at 0.965, Inception at 0.965, and MobileNet at 0.9528. Furthermore, topping ResNet at 0.9570, Inception at 0.9501, and MobileNet at 0.9311 was the ROC AUC score, 0.9698. Overall, GlaucoNet shows a satisfactory performance.

Confusion matrix study confirmed GlaucoNet's performance by displaying among the compared models the best precision (0.9706) and recall (0.9732). These results verify that combining several architectures enhances model dependability and correctness. GlaucoNet showed high diagnostic accuracy on public datasets; further validation on clinical data is needed before clinical application.

The inclusion of CIs and AUC-based statistical testing (DeLong test) strengthens the reliability of the reported metrics and confirms that the observed improvement of GlaucoNet over baseline CNNs is statistically significant rather than because of random variation.

The proposed GlaucoNet framework demonstrated high diagnostic accuracy and robustness across two benchmark public datasets, ORIGA and ACRIMA. These findings suggest strong potential for automated glaucoma detection. However, the current results are limited to retrospective datasets. Therefore, prospective clinical validation, comparison with expert ophthalmologist grading, and real-world deployment studies are required before GlaucoNet can be considered a reliable clinical screening tool for early-stage glaucoma.

Future work will focus on extending the GlaucoNet framework beyond its current technical evaluation to achieve stronger clinical, ethical, and practical integration. A key objective will be to validate the model using larger, diverse, and hospital-acquired datasets, including prospective studies involving ophthalmologist assessments, to strengthen clinical correlation and generalisability. Advanced validation methods such as *k*-fold cross-validation and multicentre external testing will be employed to ensure robust and reproducible performance. Further exploration of adaptive ensemble strategies, weighted fusion, and explainable AI techniques such as Grad-CAM will enhance interpretability, allowing clinicians to visualise decision regions and build trust in

AI-assisted screening. In addition, ethical compliance and adherence to dataset licensing norms will be maintained, with future collaborations aimed at obtaining institutional ethical approvals for real-world data collection. The regulatory aspects concerning medical AI deployment – such as conformity with FDA and CE standards – will also be studied to align the model with clinical safety requirements. Overall, future research will transform GlaucoNet from a high-performing experimental model into a clinically validated, ethically compliant, and interpretable AI-based diagnostic support tool for glaucoma screening and early detection. External clinical validation on unseen institutional datasets will be performed in future work to test the cross-dataset generalisability of GlaucoNet. Future studies will focus on validating GlaucoNet using independent institutional datasets and conducting clinician-AI comparative assessments to evaluate diagnostic consistency and interpretability. This will be essential to establish the model's true clinical reliability and regulatory compliance for ophthalmic applications.

Acknowledgements

Competing interests

The authors declare that they have no financial or personal relationships that may have inappropriately influenced them in writing this article.

CRediT authorship contribution

Minakshi Kumar: Conceptualisation, Data curation, Formal analysis, Methodology, Writing – original draft. Deepak Dembla: Conceptualisation, Formal analysis, Investigation, Methodology, Supervision, Writing – original draft. Vishal Goyal: Resources, Supervision, Writing – original draft, Writing – review & editing. All authors reviewed the article, contributed to the discussion of results, approved the final version for submission and publication, and take responsibility for the integrity of its findings.

Funding information

The authors received no financial support for the research, authorship, and/or publication of this article.

Data availability

The authors confirm that the data supporting the findings of this study are available within the article and its references.

Disclaimer

The views and opinions expressed in this article are those of the authors and are the product of professional research. They do not necessarily reflect the official policy or position of any affiliated institution, funder, agency, or that of the publisher. The authors are responsible for this article's results, findings, and content.

References

1. Quigley HA, Broman AT. The number of people with glaucoma worldwide in 2010 and 2020. Br J Ophthalmol. 2006;90(3):262–267. https://doi.org/10.1136/bjo.2005.081224
2. Weinreb RN, Aung T, Medeiros FA. The pathophysiology and treatment of glaucoma: A review. JAMA. 2014;311(18):1901–1911. https://doi.org/10.1001/jama.2014.3192
3. Ting DSW, Pasquale LR, Peng L, et al. Artificial intelligence and deep learning in ophthalmology. Br J Ophthalmol. 2019;103(2):167–175. https://doi.org/10.1136/bjophthalmol-2018-313173
4. Lee YH, Lee MM, De Silva DM, et al. Autocrine signaling by receptor tyrosine kinases in urothelial carcinoma of the bladder. PLoS One. 2020;15(10):e0241766. https://doi.org/10.1371/journal.pone.0241766
5. Li Z, He Y, Keel S, Meng W, Chang RT, He M. Efficacy of a deep learning system for detecting glaucomatous optic neuropathy based on color fundus photographs. Ophthalmology. 2018;125(8):1199–1206. https://doi.org/10.1016/j.ophtha.2018.01.023
6. Hu X, Zhang L-X, Gao L, et al. GLIM-Net: Chronic glaucoma forecast transformer for irregularly sampled sequential fundus images. IEEE Trans Med Imaging. 2023;42(6):1875–1884. https://doi.org/10.1109/TMI.2023.3243692
7. Srinivasan PP, Kim LA, Mettu PS, Cousins SW. Fundus photography: Principles, applications and future directions. Retina Today. 2015;10(6):28–32.
8. Jonas JB, Budde WM, Panda-Jonas S. Ophthalmoscopic evaluation of the optic nerve head. Surv Ophthalmol. 1999;43(4):293–320. https://doi.org/10.1016/S0039-6257(99)00009-7
9. Burgoyne CF. A biomechanical paradigm for axonal insult within the optic nerve head in aging and glaucoma. Exp Eye Res. 2011;93(2):120–132. https://doi.org/10.1016/j.exer.2010.09.005
10. Tham YC, Li X, Wong TY, Quigley HA, Aung T, Cheng CY. Global prevalence of glaucoma and projections of glaucoma burden through 2040: A systematic review and meta-analysis. Ophthalmology. 2014;121(11):2081–2090. https://doi.org/10.1016/j.ophtha.2014.05.013
11. Muhammad H, Fraz MM, Remagnino P, et al. An ensemble classification-based approach applied to retinal nerve fiber layer defect detection in fundus images. Comput Methods Programs Biomed. 2017;137:231–241. https://doi.org/10.1016/j.cmpb.2016.09.013
12. Raghavendra U, Fujita H, Gudigar A, et al. Deep convolution neural network for accurate diagnosis of glaucoma using digital fundus images. Inf Sci. 2018;441:41–49. https://doi.org/10.1016/j.ins.2018.02.005
13. Moon CH, Hwang YH, Park K, Kim Y. Predicting 10-2 visual fields from wide-field swept-source OCT using deep learning. Ophthalmol Sci. 2021;1(2):100059. https://doi.org/10.1016/j.xops.2021.100059
14. Zarzà M, Carrera A, Pardo A. Glaucoma detection using efficient CNN architectures and novel training strategies. Sensors. 2021;21(17):5701. https://doi.org/10.3390/s21175701
15. Park K, Kim YK, Park KH, Jeoung JW. Comparison of deep learning models for visual field prediction in glaucoma. Transl Vis Sci Technol. 2021;10(4):18. https://doi.org/10.1167/tvst.10.4.18
16. Tong Y, Liu Y, Zhao M, Meng L, Zhang J. Improved U-net MALF model for lesion segmentation in breast ultrasound images. Biomed Signal Process Control. 2021;68:102721. https://doi.org/10.1016/j.bspc.2021.102721
17. Gupta V, Wang H, Wen JC. AI-based early glaucoma detection using confocal scanning laser ophthalmoscopy. J Glaucoma. 2020;29(2):89–95. https://doi.org/10.1097/IJG.0000000000001412
18. Zhang Y, Li S, Wang X, et al. Multimodal deep learning models for glaucoma diagnosis using OCT and visual fields. IEEE Access. 2020;8:103575–103584. https://doi.org/10.1109/ACCESS.2020.2999118
19. Islamaj R, Wei C-H, Cissel D, et al. NLM-Gene, a richly annotated gold standard dataset for gene entities that addresses ambiguity and multi-species gene recognition. J Biomed Inform. 2021;118:103779. https://doi.org/10.1016/j.jbi.2021.103779
20. Guo Y, Chen H, Zhang X, et al. Development and validation of a mobile app for glaucoma diagnosis using AI. JMIR mHealth uHealth. 2021;9(3):e24454. https://doi.org/10.2196/24454
21. Hemelings R, Elen B, Barbosa-Breda J, et al. Transfer and active learning for glaucoma detection with deep neural networks. Comput Med Imaging Graph. 2020;84:101720. https://doi.org/10.1016/j.compmedimag.2020.101720
22. Long D, Thompson A, Vermeer KA, et al. Simplifying machine learning models for real-world glaucoma prediction. PLoS One. 2021;16(2):e0245613. https://doi.org/10.1371/journal.pone.0245613
23. Kumari A, Pandey S, Rajalakshmi R. Ensemble learning-based glaucoma detection in fundus images. Expert Syst Appl. 2022;189:116081. https://doi.org/10.1016/j.eswa.2021.116081
24. Sharifi M, Khatibi T, Emamian MH, Sadat S, Hashemi H, Fotouhi A. Development of glaucoma predictive model and risk factors assessment based on supervised models. BioData Min. 2021;14(1):48.
25. Diaz-Pinto A, Morales S, Naranjo V, Köhler T, Mossi JM, Navea A. CNNs for automatic glaucoma assessment using fundus images: an extensive validation. Biomed Eng Online. 2019;18(1):29.
26. Bremond-Gignac D, Dairazalia SC, Jihyun LE, et al. Preventive screening of open-angle glaucoma: an innovative machine learning risk assessment tool based on health insurance claims data. Acta Ophthalmol. 2022;100. https://doi.org/10.1111/j.1755-3768.2022.089

27. Li L, Xu M, Liu H, Zhou Y, Dai X, Wang Z. A large-scale database and a CNN model for attention-based glaucoma detection. IEEE Trans Med Imaging. 2019;38(11): 2468–2478. https:doi.org/10.1109/TMI.2019.2918417
28. Gupta K, Thakur A, Goldbaum M, Yousefi S. Glaucoma precognition: Recognizing preclinical visual functional signs of glaucoma. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops; 2020 June 14–19; Seattle, WA, USA, Piscataway, NJ, USA: IEEE/CVF; 2020. p. 1020–1021.
29. Hu B, Zhou S, Xiong Z, Wu F. Recursive decomposition network for deformable image registration. IEEE J Biomed Health Inform. 2022;26(10):5130–5141.
30. Lee Y, Kim YS, Lee DI, Jeong S, Kang GH, Jang YS, et al. The application of a deep learning system developed to reduce the time for RT-PCR in COVID-19 detection. Sci Rep. 2022;12(1):1234.
31. Juneja M, Singh S, Agarwal N, et al. Automated detection of Glaucoma using deep learning convolution network (G-net). Multimed Tools Appl. 2020;79(21):15531–15553.
32. Kamal MS, Dey N, Chowdhury L, Hasan SI, Santosh KC. Explainable AI for glaucoma prediction analysis to understand risk factors in treatment planning. IEEE Trans Instrum Meas. 2022;71:1–9.
33. Akter N, Fletcher J, Perry S, Simunovic MP, Briggs N, Roy M. Glaucoma diagnosis using multi-feature analysis and a deep learning technique. Sci Rep. 2022;12(1):8064.
34. Lee EJ, Kim TW, Kim JA, Lee SH, Kim H. Predictive modeling of long-term glaucoma progression based on initial ophthalmic data and optic nerve head characteristics. Transl Vis Sci Technol. 2022;11(10):24.
35. Nakahara K, Asaoka R, Tanito M, et al. Deep learning-assisted (automatic) diagnosis of glaucoma using a smartphone. Br J Ophthalmol. 2022;106(4):587–592.
36. Gao XR, Huang H, Kim H. Polygenic risk score is associated with intraocular pressure and improves glaucoma prediction in the UK Biobank cohort. Transl Vis Sci Technol. 2019;8(2):10.
37. Devecioglu OC, Malik J, Ince T, Kiranyaz S, Atalay E, Gabbouj M. Real-time glaucoma detection from digital fundus images using self-onns. IEEE Access. 2021;9:140031–41.
38. Jha P, Dembla D, Dubey W. Implementation of Machine Learning Classification Algorithm Based on Ensemble Learning for Detection of Vegetable Crops Disease. Int J Adv Comput Sci Appl. 2024;15(1)
39. Jha P, Dembla D, Dubey W. Implementation of transfer learning based ensemble model using image processing for detection of potato and bell pepper leaf diseases. Int J Intell Syst Appl Eng. 2024;12(8s):69–80
40. Gampala V, Maram B, Vigneshwari S, Cristin R. Glaucoma detection using hybrid architecture based on optimal deep neuro fuzzy network. Int J Intell Syst. 2022;37(9):6305–30.
41. Lavric A, Petrariu AI, Havriliuc S, Coca E. Glaucoma detection by artificial intelligence: GlauNet a deep learning framework. 2021 International Conference on e-Health and Bioengineering (EHB); 2021 Oct 14–16; Iasi, Romania, Piscataway, NJ, USA: Institute of Electrical and Electronics Engineers (IEEE); 2021. p. 1–4.
42. Meedeniya D, Shyamalee T, Lim G, Yogarajah P. Glaucoma identification with retinal fundus images using deep learning: Systematic review. Inform Med Unlocked. 2025;56:101644. https://doi.org/10.1016/j.imu.2025.101644
43. Diaz-Pinto A, Morales S, Naranjo V, Köhler T, Mossi JM, Navea A. CNNs for automatic glaucoma assessment using fundus images: an extensive validation. Biomedical engineering online. 2019 Mar 20;18(1):29.
44. Islam MZ, Hossain MS, ul Islam R, Andersson K. Static hand gesture recognition using convolutional neural network with data augmentation. 2019 Joint 8th International Conference on Informatics, Electronics & Vision (ICIEV) and 2019 3rd International Conference on Imaging, Vision & Pattern Recognition (icIVPR); 2019 May 30–Jun 2; Spokane, WA, USA, Piscataway, NJ, USA: IEEE; 2019. p. 324–329.
45. Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. Commun ACM. 2012;60(6):84–90. https://doi.org/10.1145/3065386
46. JCivit-Masot J, Domínguez-Morales MJ, Vicente-Díaz S, Civit A. Dual machine-learning system to aid glaucoma diagnosis using disc and cup feature extraction. IEEE Access. 2020;8:127519–127529.
47. Gómez-Valverde JJ, Antón A, Fatti G, et al. Automatic glaucoma classification using color fundus images based on convolutional neural networks and transfer learning. Biomed Opt Express. 2019;10(2):892–913.
48. Serener A, Serte S. Transfer learning for early and advanced glaucoma detection with convolutional neural networks. In: 2019 Medical technologies congress (TIPTEKNO); 2019 Oct 3; IEEE, 2019; p. 1–4.
49. Jiang Y, Duan L, Cheng J, et al. JointRCNN: a region-based convolutional neural network for optic disc and cup segmentation. IEEE Trans Biomed Eng. 2019;67(2): 335–343.