STM JOURNALS

International Journal of
Brain Sciences

ISSN: 3048-7935
Volume 2, Issue 2, 2025
July–December
DOI (Journal): 10.37591/IJBS

https://journals.stmjournals.com/ijbs

Research IJBS

Enhancing Glaucoma Diagnosis with Deep Learning: A Study Using ResNet-50 and DenseNet-121

Meenakshi Sharma[1,*], Naman Sharma[2], Pinu Adarsh[2], Vikash Prasad[2]

Abstract

Glaucoma is a leading cause of irreversible blindness worldwide, mainly resulting from progressive optic nerve damage, often related to elevated intraocular pressure. Early detection is essential to prevent vision loss, but traditional diagnostic methods rely on specialized equipment and trained professionals, making large-scale screening difficult. This study uses a publicly available fundus imaging dataset to explore the effectiveness of deep learning models for glaucoma detection. These datasets provide medical images, specifically fundus images (images of the back of the eye), and specifically evaluate the performance of ResNet-50 and DenseNet-121 architectures in binary and multi-class classification tasks. In a binary classification task for glaucoma detection, the combined use of ResNet-50 and DenseNet-121 models delivered impressive results – achieving a precision of 1.0, recall of 0.92, specificity of 1.0, an F1 score of 0.958, and an overall accuracy of 96.1%. These findings suggest that binary classification offers better performance than multi-class approaches for identifying glaucoma. The study emphasizes the promise of deep learning techniques in supporting early glaucoma diagnosis and highlights the importance of continued clinical validation before these tools can be widely adopted in real-world settings.

Keywords: Glaucoma, Deep Learning, ResNet-50, DenseNet-121, medical image analysis

INTRODUCTION

Glaucoma is a progressive and often silent eye condition that gradually damages the optic nerve, commonly linked to elevated intraocular pressure (IOP) [1]. It's one of the leading causes of permanent blindness across the globe, affecting millions of individuals. As reported by the World Glaucoma Association, around 80 million people were living with glaucoma in 2022, and that number is projected to rise to 111.8 million by 2040 [2]. Because the disease usually shows no symptoms in its early stages, many people are only diagnosed after considerable vision loss has already occurred. Detecting glaucoma early is critical to slowing its progression and beginning timely treatment. Standard diagnostic methods – such as visual field analysis, optical coherence tomography (OCT), and tonometry – are widely used and reliable. However, these approaches depend on expensive equipment and skilled professionals, which can limit their availability for large-scale screening efforts. Moreover, these methods are prone to human error, further emphasizing the need for automated and AI-based diagnostic solutions [3].

Medical image segmentation has become a crucial technique in glaucoma detection, enabling accurate identification of affected regions in retinal images. Over the past two decades, deep learning (DL) techniques have significantly advanced medical image analysis, particularly in

*Author for Correspondence
Meenakshi Sharma
E-mail: shrama.minaxi@gmail.com

[1]Professor, Department of Computer Science & Engineering, Global Group of Institutes, Amritsar, Punjab, India
[2]Student, Professor, Department of Computer Science & Engineering, Global Group of Institutes, Amritsar, Punjab, India

Received Date: April 25, 2025
Accepted Date: July 09, 2025
Published Date: July 14, 2025

Citation: Meenakshi Sharma, Naman Sharma, Pinu Adarsh, Vikash Prasad. Enhancing Glaucoma Diagnosis with Deep Learning: A Study Using ResNet-50 and DenseNet-121. International Journal of Brain Sciences. 2025; 2(2): 9–18p.

segmentation and classification tasks [4]. Convolutional neural networks (CNNs) have shown great promise in glaucoma detection, with fully convolutional networks (FCNs), such as U-Net being widely used for retinal image segmentation. Traditional U-Net architectures use an encoder-decoder structure to extract features and reconstruct images, while Attention U-Net improves segmentation accuracy by focusing on key target structures of various shapes and sizes [5].

However, conventional U-Net and Attention U-Net architectures lack batch normalization (BN), which can lead to overfitting and suboptimal performance in medical image analysis [6]. To address these limitations, this study proposes an improved approach by integrating Attention U-Net with CNN architectures for fundus image segmentation followed by classification for glaucoma detection. This combined method improves feature extraction, enhances classification accuracy, and contributes to the development of reliable AI-based diagnostic tools for glaucoma screening. In recent years, CNN-based models, such as InceptionV3, VGG19, and ResNet-50 have shown remarkable success in medical image analysis. In this study, we fine-tuned these models for segmentation to enhance feature extraction and classification accuracy.

InceptionV3 efficiently utilizes computing resources and extracts features at multiple scales using different convolutional filter sizes. VGG19, a deep 19-layer CNN, follows a structured architecture of convolutional and max-pooling layers, concluding with fully connected layers and a SoftMax classifier. ResNet-50 addresses the vanishing gradient problem using identity mapping, allowing deeper networks to train effectively without increasing the error rate [7]. By integrating these architectures with Attention U-Net, our approach enhances fundus image segmentation and glaucoma detection, improving both accuracy and reliability. The rest of this paper is organized as follows: Section II provides an overview of previous research related to glaucoma detection using deep learning techniques. Section III explains the methodology, including a detailed description of the model architecture and the data set employed. Section IV presents the experimental findings and evaluates the performance of the proposed method. Lastly, Section V summarizes the study’s key insights and suggests possible directions for future research.

RELATED WORK

The application of CNNs in glaucoma detection has been widely explored, with significant advances in recent years. Early studies mainly focused on binary classification, where CNN-based models demonstrated superior performance in glaucoma detection from fundus images compared to traditional classifiers. These studies have played a crucial role in paving the way for ongoing advancements in glaucoma diagnosis using deep learning techniques [8]. As research progressed, multi-class classification approaches were introduced, enabling CNNs to distinguish different stages of glaucoma. Models, such as InceptionV3, VGG19, and ResNet-50 have shown remarkable performance in medical image analysis, including fundus image segmentation and classification. InceptionV3 efficiently extracts multi-scale features, VGG19 uses a structured deep architecture, and ResNet-50 alleviates the vanishing gradient problem, enabling the efficient training of deeper networks [9].

In addition, fully convolutional networks (FCNs), such as U-Net and Attention U-Net have attracted attention for medical image segmentation. While U-Net follows an encoder-decoder structure for accurate segmentation, Attention U-Net improves feature selection by focusing on critical regions of the optic nerve. These architectures have been instrumental in improving automated glaucoma detection [10]. Recent advances have explored hybrid models that combine CNN-based classifiers with segmentation networks to improve feature extraction, classification accuracy, and model interpretability. By integrating Attention U-Net with CNN architectures, the researchers achieved improved segmentation performance, contributing to more reliable AI-based glaucoma screening methods. Binary classification has become a popular choice for detecting glaucoma, thanks to its straightforward approach and strong diagnostic performance. CNN-based binary classifiers have shown excellent sensitivity and specificity, enhancing the effectiveness of deep learning in this field [11]. To further improve the classification performance, ensemble methods combining multiple deep CNN architectures have been introduced, demonstrating improved accuracy [12]. Lightweight CNN ensembles optimized for glaucoma detection on mobile have also been developed, providing a practical solution for low-resource environments [13]. In addition, transfer learning using pre-trained convolutional neural networks (CNNs) has shown great promise, enabling models to perform well even when training data is limited. Popular architectures, like ResNet-50 and DenseNet-121, have become key tools in glaucoma detection. ResNet-50 is well regarded for its cutting-edge performance in glaucoma classification, thanks to its use of residual connections that help streamline and enhance the training process [14].

DenseNet-121, with its ability to capture complex retinal features, has shown effectiveness in medical imaging tasks, further advancing deep learning applications in glaucoma detection [15]. Recent research has investigated cross-modal learning approaches, where models are trained using various types of imaging data. One promising method involves a multi-task learning framework that combines fundus photographs with OCT scans. By integrating these complementary imaging modalities, the approach has shown enhanced performance in glaucoma detection [16].

This approach is beneficial in low-resource settings, where large, annotated datasets may not be available, making AI- driven glaucoma detection more accessible and efficient.

METHODOLOGY

The proposed deep learning-based model for glaucoma detection integrates multiple steps including pre-processing, data augmentation, segmentation, classification, and evaluation as shown in Figure 1. The framework uses fundus images as input, which undergo pre-processing techniques, such as contrast-limited adaptive equalization (CLAHE) and dilation to improve image quality and highlight critical features.

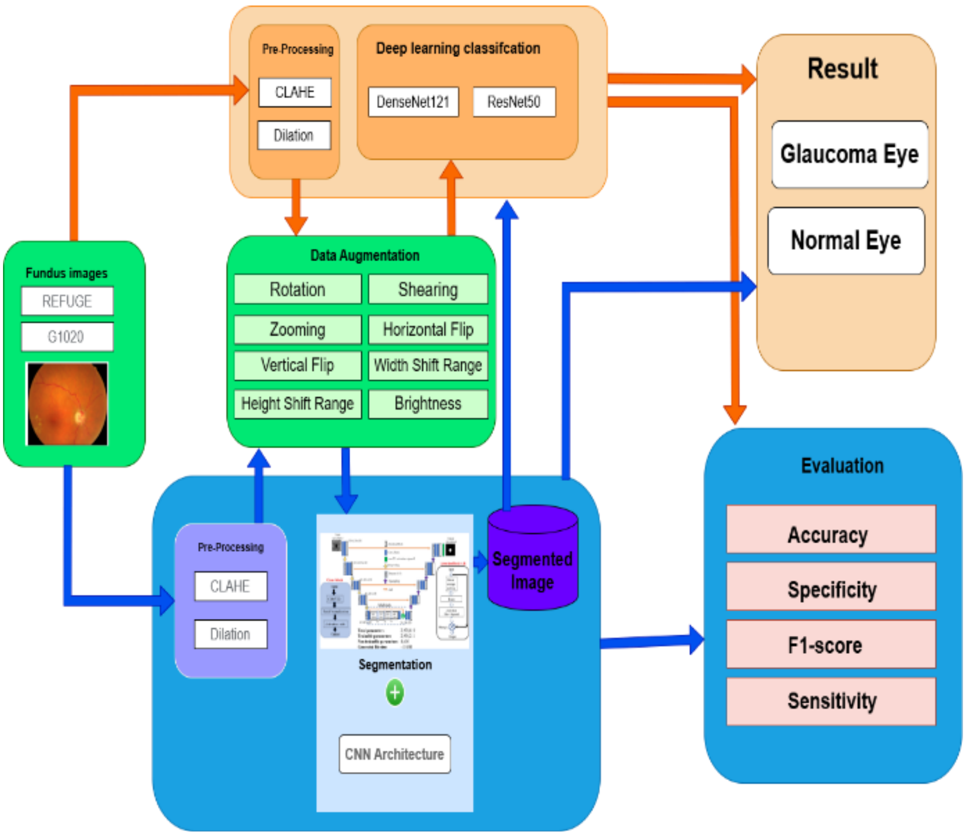

Figure 1. Proposed framework model.

These pre-processed images are then subjected to data augmentation methods including rotation, shearing, flipping, zooming, brightness adjustment, and shift range variations. This augmentation strategy improves the generalization ability of the model by simulating real-world variations in retinal images. A segmentation module based on a convolutional neural network (CNN) is used to extract relevant structural features from fundus images, generating segmented images that preserve essential information for classification. The classification step leverages two state-of-the-art deep learning architectures, DenseNet-121 and ResNet-50, to perform robust feature extraction and classification. Together, these models improve predictive performance by combining their strengths in hierarchical feature learning.

The final classification output determines whether the input image corresponds to a "normal eye" or a "glaucomatous eye". To ensure a comprehensive evaluation, the model performance is assessed using several metrics, including accuracy, specificity, sensitivity, and F1 score. These evaluation metrics provide insight into the model's effectiveness in distinguishing glaucomatous and non-glaucomatous cases, thereby, ensuring its reliability in clinical applications.

By integrating advanced pre-processing, data augmentation, segmentation, and deep learning-based classification, the proposed model offers a highly efficient and automated approach for glaucoma detection. This framework has the potential to assist ophthalmologists in early diagnosis and treatment planning, ultimately contributing to improved patient outcomes in glaucoma management. This study presents a structured framework for glaucoma detection using deep learning. The process starts with data preparation, including loading and augmentation of fundus images to improve the diversity of the datasets. After pre-processing, features are extracted using pre-trained convolutional neural networks (CNNs), specifically ResNet-50 and DenseNet-121. To boost performance, an ensemble method is applied to combine the strengths of both models. This ensemble strategy helps enhance classification accuracy. The training process is optimized using the Adam optimizer, and the model's effectiveness is assessed through important evaluation metrics, including accuracy, sensitivity, specificity, and the F1 score.

The overall framework, shown in Figure 2, demonstrates the effectiveness of the ensemble approach in glaucoma classification.

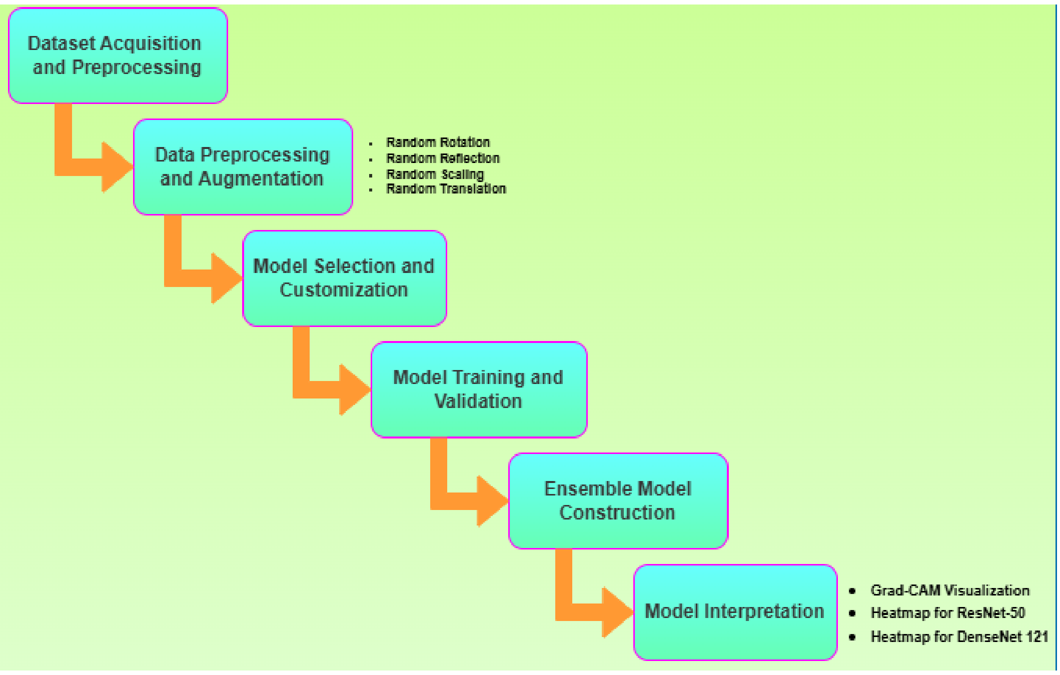

Figure 2. Framework AI-driven glaucoma detection.

Data Preparation

The dataset, consisting of fundus images, was processed using Jupyter Notebook with Google Colab. The images were carefully loaded and managed using Python tools like TensorFlow and PyTorch. The dataset was then split into training and testing groups, with 80% of the images used for training the model and the remaining 20% reserved for testing its performance.

Data Augmentation

To help the model perform better on new, unseen data and avoid overfitting, various data augmentation methods were applied to the training images. These techniques introduced variations into the dataset, ensuring that the model could recognize glaucomatous patterns under different conditions. Figure 3 illustrates the augmentation methods, which include:

- *Random rotations (–30° to 30°)*: Fundus images were rotated within this range to account for variations in patient positioning and camera alignment. This helped the model learn to detect glaucoma features regardless of orientation.
- *Random horizontal and vertical flip*: To simulate real-world imaging variations, images were flipped horizontally and vertically. This prevented the model from developing biases toward specific orientations and improved its ability to recognize glaucoma patterns in various scenarios.
- By leveraging Jupyter Notebook with Google Colab for the implementation and application of these pre-processing techniques, the proposed framework improves the model performance in binary and multi-class glaucoma classification tasks.

Model Selection and Modification

For this study, we selected two pre-trained CNN models: ResNet-50 and DenseNet-121. These architectures were chosen due to their strong track record in image classification. To adapt them to our specific task, we fine-tuned each model by replacing the final fully connected layer and classification layer to fit the number of classes in our dataset.

Model Training

The updated models were trained using the Adam optimizer, chosen for its efficiency and its ability to effectively manage sparse gradients. Adam combines the strengths of AdaGrad, which works well with sparse data, and RMSProp, known for its success in online learning. By adjusting the learning rate for each parameter individually, Adam helps the model converge more quickly and steadily – an important factor when working with complex medical images. Training was carried out over 15 epochs with a mini-batch size of 32, starting with a learning rate of 0.0001. To keep track of the model's progress, validation was performed on the test set every 30 iterations throughout training.

Ensemble Model Prediction

To enhance the accuracy of classification, we combined the predictions from ResNet-50 and DenseNet-121 by averaging their outputs to arrive at the final decision. The predicted class label for each image was determined based on the highest average probability across both models. This ensemble strategy leveraged the strengths of both architectures, improving overall performance in glaucoma classification.

Model Evaluation

Model performance was evaluated using standard evaluation metrics to ensure a comprehensive and objective analysis. The following key metrics were used for both binary and multi-class classification:

- *Precision:* Measures the overall accuracy of the model predictions.
- *Sensitivity (Recall):* Measures how well the model can correctly detect glaucoma cases.
- *Specificity:* Assesses the model's ability to distinguish between healthy and glaucoma-affected images.
- *F1-score:* Strikes a balance between precision and recall, making it especially valuable when dealing with imbalanced datasets.

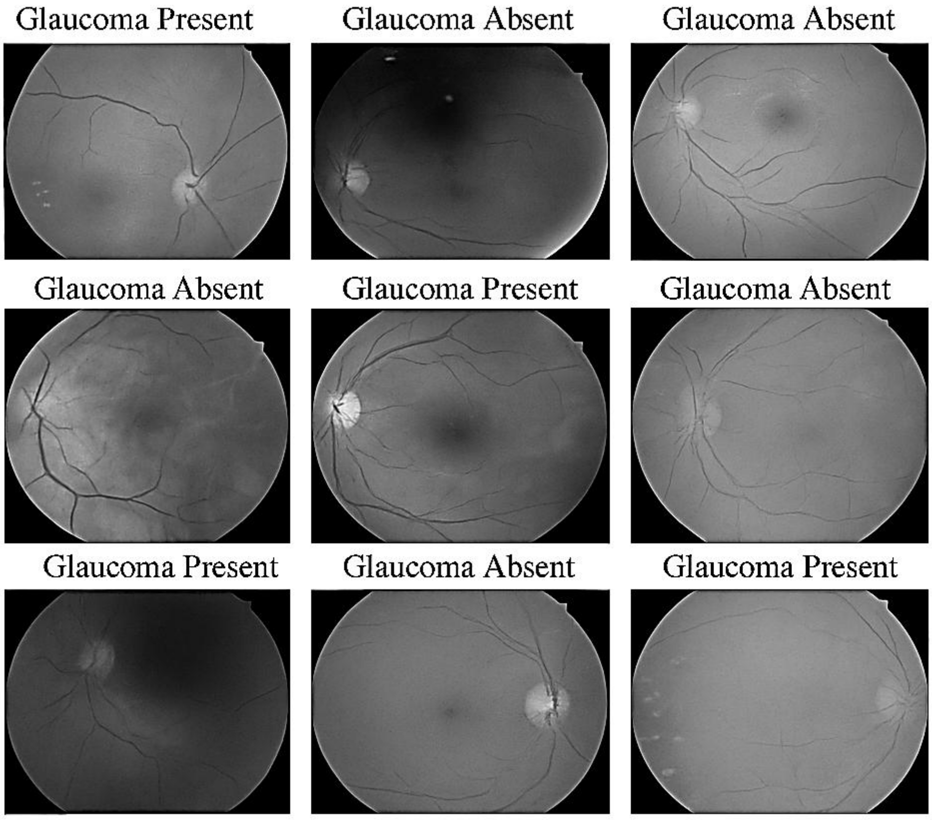

Figure 3. Data augmentation of fundus images.

These metrics helped users assess how well the models could classify images into normal, early-stage glaucoma, and advanced glaucoma categories. The evaluation was conducted systematically to ensure the reliability and robustness of the results.

RESULT AND DISCUSSION

The dataset titled Glaucoma Fundus Imaging Data-sets compiles fundus images along with optic disc (OD) and optic cup (OC) masks from three major sources:

1. ORIGA.
2. REFUGE.
3. G1020.

In total, these datasets comprise 2,870 fundus images.

Key Components

- *ORIGA Dataset:* Sourced from the Singapore National Eye Centre, the ORIGA dataset comprises 650 annotated fundus images, with detailed segmentations of the OD and OC.
- *REFUGE Dataset:* Part of the Retinal Fundus Glaucoma Challenge, the REFUGE dataset comprises 1,200 fundus images, each with OD and OC annotations, facilitating a comprehensive glaucoma analysis.
- *G1020 Dataset:* This dataset consists of 1,020 fundus images with corresponding OD and OC masks, providing a substantial resource for training and evaluating glaucoma detection models.

For binary classification, the Early Glaucoma and Advanced Glaucoma categories were combined into a single class "Glaucomatous," resulting in two final categories: Normal and Glaucomatous. While the dataset provides useful information, its small size makes it challenging to guarantee that the model will perform well on new data from diverse populations. Retinal images can differ significantly due to factors, like age, ethnicity, the type of imaging equipment, and various health conditions – many of which might not be fully represented in the current dataset. Because of this limited sample size, there's a risk the model could overfit, doing well on the training images but struggling with new, unseen data. To help prevent this, techniques, like random rotations, flips, and scaling were used to artificially expand and diversify the training data. These methods improved the robustness of the model by introducing variations that simulate real-world scenarios. However, while augmentation improves performance, it cannot completely replace a larger, more diverse dataset that naturally captures these variations.

Evaluation Metrics

The effectiveness of both the binary and multi-class classification models was assessed using common metrics based on the confusion matrix. For binary classification, accuracy was calculated using the following formula: it represents the proportion of true positive cases (TP) out of all positive predictions made by the model. In other words, it shows how many images labeled as glaucomatous were correctly identified.

False positives (FP) occur when the model mistakenly labels a healthy eye as having glaucoma. To evaluate how accurately the model identifies actual glaucoma cases, accuracy was calculated using Equation (1), offering a clear measure of its overall performance.

$$Precision = \frac{\text{True Posirive}}{\text{True Positive+False Posiive}} \quad (1)$$

Recall measures the model's ability to correctly identify positive cases using Equation (2), it is calculated as the ratio of true positives (TP) to the total number of true positive cases, which includes both true positives (TP) and false negatives (FN).

$$Recall = \frac{\text{True Posirive}}{\text{True Positive+False Negative}} \quad (2)$$

A false negative (FN) happens when the model fails to identify a real glaucoma case, mistakenly classifying it as normal. When recall is high, it means the model is correctly detecting most actual glaucoma cases, which helps reduce the number of false negatives.

Performance Evaluation

To evaluate the effectiveness of glaucoma classification, key performance indicators, such as accuracy, specificity, and sensitivity, were analyzed. Sensitivity and specificity are crucial quantitative measures that indicate the reliability of the classification model in distinguishing between glaucomatous and non-glaucomatous cases. These indicators help determine the ability of the model to correctly detect the presence or absence of the disease. Calculations of these parameters are performed using Equations (3) and (4), respectively.

$$Specificity = \frac{\text{True Negative}}{\text{True Negative+False Posiive}} \quad (3)$$

$$Accuracy = \frac{\text{TP+TN}}{\text{TP+TN+FP+FN}} \quad (4)$$

The F1 score combines both precision and recall into a single metric by calculating their weighted average, considering true positives and false negatives. While precision alone is often used to evaluate models, the F1 score is especially helpful when working with imbalanced datasets, as it offers a more balanced and meaningful measure of performance. Equation (5) was used to compute the F1 score.

$$\text{F1 Score} = \frac{\text{2 X recision X Recall}}{\text{Precision+Recall}} \quad (5)$$

The highest classification accuracy is achieved when the costs of false positives and false negatives are balanced, ensuring a well-optimized and reliable model.

Binary Classification Results

For the binary classification task, the dataset was divided into two categories: "Normal" and "Glaucomatous", with the "Glaucomatous" class encompassing both "Early Glaucoma" and "Advanced Glaucoma" images. The ensemble model, incorporating the predictions of ResNet-50 and DenseNet-201, showed good performance in distinguishing these classes.

Comparison of Optimizer Performance on Different Batch Sizes

The impact of different optimizers and batch sizes on classification performance was analyzed as shown in Tables 1 and 2. The Adam optimizer consistently achieved high accuracy across all batch sizes, with the highest performance of 98.3% for a batch size of 64. The RMSprop optimizer showed stable performance, reaching a maximum accuracy of 98.1%, while the SGD optimizer demonstrated gradual improvement, reaching 98.2% for a batch size of 64.

Table 1. Performance classification of ResNet-50.

Optimizer	Batch Size 8	Batch Size 16	Batch Size 32	Batch Size 64
Adam	97.8%	98.1%	97.9%	98.2%
RMSprop	97.6%	97.9%	98.0%	98.1%
SGD	97.5%	97.8%	98.0%	98.1%

Table 2. Performance classification of DenseNet.

Optimizer	Batch Size 8	Batch Size 16	Batch Size 32	Batch Size 64
Adam	97.8%	98.0%	98.1%	98.3%
RMSprop	97.9%	98.0%	98.1%	98.1%
SGD	97.9%	97.9%	98.0%	98.2%

Overall, the results indicate that Adam outperformed RMSprop and SGD, especially for larger batch sizes, making it the most effective optimizer for this classification task. For ResNet-50, the Adam optimizer demonstrated the highest accuracy for different batch sizes as shown in Figure 4, especially for batch size 64, where it reached about 98.2%. RMSprop and SGD followed a similar trend, with SGD showing the lowest accuracy for smaller batch sizes but gradually improving as the batch size increased.

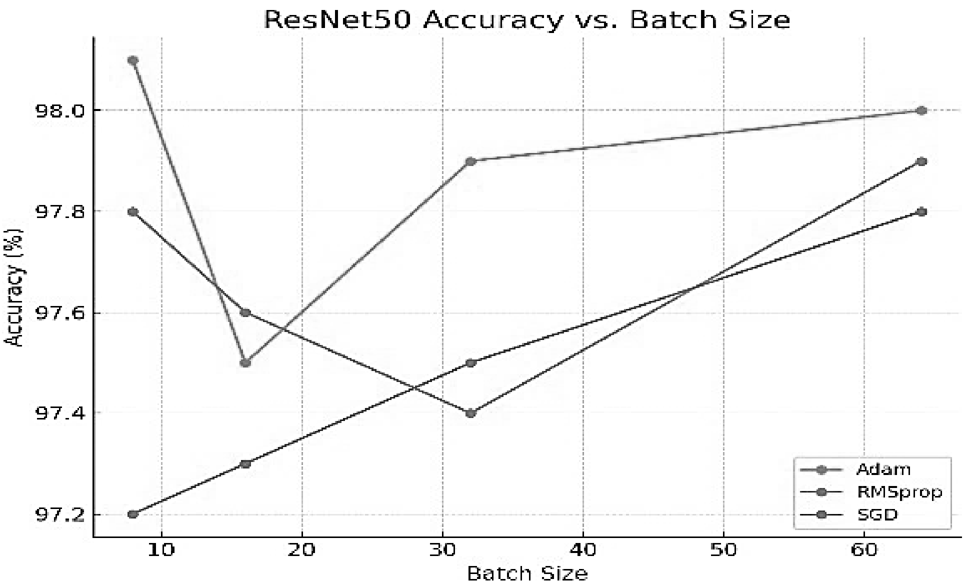

Figure 4. ResNet-50 accuracy vs batch size.

Similarly, for DenseNet-121, Adam consistently outperformed the other optimizers, reaching its highest accuracy of about 98.4% for batch size 64 shown in Figure 5, RMSprop and SGD showed an increasing trend in accuracy as the batch size increased, with SGD showing the most significant improvement for different batch sizes.

Overall, Adam was found to be the most efficient optimizer for both models, especially for larger batch sizes, suggesting its superiority in optimizing deep learning models for glaucoma detection.

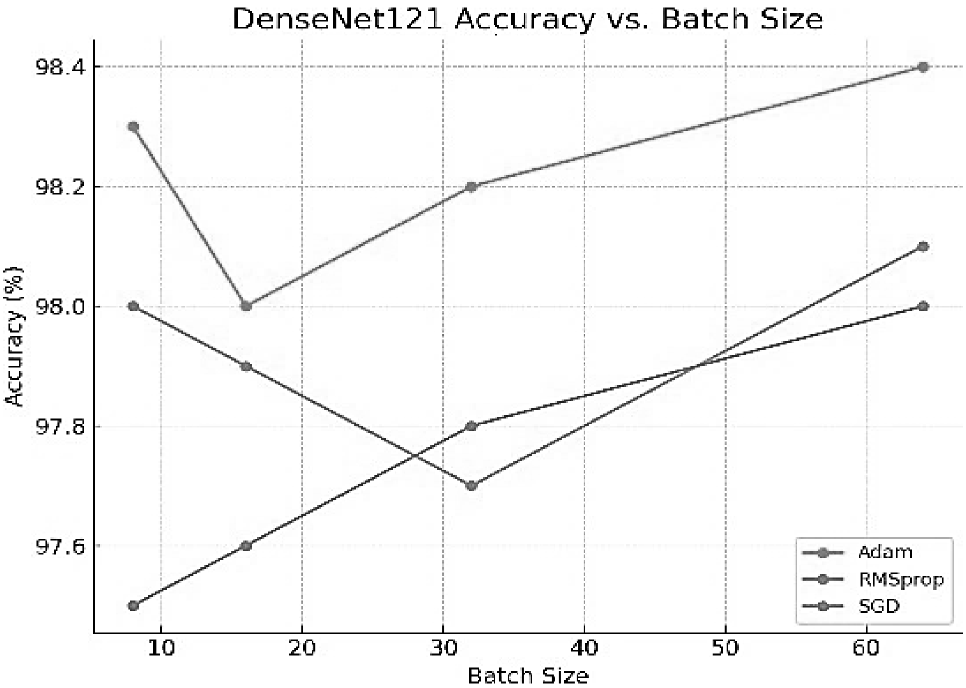

Figure 5. DenseNet-121 accuracy vs batch size.

CONCLUSIONS AND FUTURE DIRECTIONS

In this study, we developed and evaluated an ensemble deep learning model for glaucoma classification using ResNet-50 and DenseNet-121 architectures. The proposed approach demonstrated high classification performance, effectively distinguishing normal and glaucomatous images. The Adam optimizer consistently outperformed RMSprop and SGD across different batch sizes, achieving the highest accuracy for larger batch sizes. Data augmentation techniques improved the robustness of the model, mitigating the effects of the limited dataset size and improving generalization. The superior performance of the ensemble model highlights its potential as a reliable tool for automated glaucoma detection, which could aid in early diagnosis and clinical decision-making. However, limitations of the study include dataset size constraints and the need for more diverse retinal images.

Future Directions

To further improve the generalization and applicability of the model in real-world clinical settings, future research will focus on:

- Expanding the dataset by incorporating images from diverse populations, different imaging devices, and diverse clinical conditions.
- Looking into more advanced deep learning models – such as Vision Transformers and hybrid CNN architectures – to enhance how features are extracted and improve the overall accuracy of classification. Integrating explainability techniques, such as Grad-CAM and SHAP to improve model interpretability and provide decision-making insights.
- Deploying the model in real-time clinical applications and validating its effectiveness in prospective studies with ophthalmologists.

By addressing these aspects, we aim to refine the proposed method, making it a more robust and clinically viable solution for automated glaucoma detection.

REFERENCES

1. World Health Organization. Report of the 2030 targets on effective coverage of eye care. Geneva: World Health Organization; 2022 Oct 12.
2. 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 Nov;121(11):2081–90.
3. Weinreb RN, Aung T, Medeiros FA. The pathophysiology and treatment of glaucoma. JAMA. 2014;311(18):1901–11.
4. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature. 2015 May 28;521(7553):436–44.
5. Prajna Y, Nath MK. Efficient blood vessel segmentation from color fundus image using deep neural network. J Intell Fuzzy Syst. 2022;42(4):3477–89.
6. Mahum R, Ur Rehman S, Okon OD, Alabrah A, Meraj T, Rauf HT. A novel hybrid approach based on deep CNN to detect glaucoma using fundus imaging. Electronics. 2022;11(1):26.
7. Ariyarathne G, De Silva S, Dayarathna S, Meedeniya D, Jayarathne S. ADHD identification using convolutional neural network with seed-based approach for MRI data. In: Proceedings of the 9th International Conference on Software and Computer Applications; 2020 Jan 9–11; Langkawi, Malaysia. New York: ACM; p. 31–5.
8. Chen X, Xu Y, Wong DWK, et al. Glaucoma detection based on deep convolutional neural networks. In: Proceedings of the 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC); 2015 Aug 25–29; Milan, Italy. IEEE; 2015.
9. 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 Jul;125(8):1199–206.
10. Christopher M, Belghith A, Bowd C, Proudfoot JA, Goldbaum MH, Weinreb RN, et al. Performance of deep learning architectures and transfer learning for detecting glaucomatous optic neuropathy in fundus photographs. Sci Rep. 2018;8:16685.
11. Kim SH, Cho J, Chae H, Lim YJ, Lee M, Oh J, et al. Efficacy of a comprehensive binary classification model using a deep convolutional neural network for wireless capsule endoscopy. Sci Rep. 2021;11:1328.
12. Ma Y, Xu T, Han S, Kim K. Ensemble learning of multiple deep CNNs using accuracy-based weighted voting for ASL recognition. Appl Sci. 2022;12:11766.
13. Sharma M, Bhagat B, Kumari V, Sharma N, Singh D. A comparative analysis of CNN and transfer learning algorithms in predicting Alzheimer's disease stages using neuroimaging. In: 2024 OPJU International Technology Conference (OTCON) on Smart Computing for Innovation and Advancement in Industry 4.0; 2024 Jun; Raigarh, India. IEEE; p. 1–6.
14. Puchaicela-Lozano MS, García-Zapirain B, Illarramendi A, Yanes J. Deep learning for glaucoma detection: R-CNN ResNet-50 and image segmentation. J Adv Inf Technol. 2024;14(6):1186–97.
15. Fu H, Cheng J, Xu Y, Zhang C, Gao N, Jin K, et al. Disc-aware ensemble network for glaucoma screening from fundus image. IEEE Trans Med Imaging. 2018 Nov;37(11):2493–501.
16. Braeu FA, Thiéry AH, Tun TA, Kadziauskiene A, Barbastathis G, Aung T, et al. Geometric deep learning to identify the critical 3D structural features of the optic nerve head for glaucoma diagnosis. Am J Ophthalmol. 2023;250:38–48.

