Emerging Trends in Digital Transformation
*Volume-1, Issue-2, July 2025* ISSN:**3107-4308** *Paper ID:* ***ETDT-V1-I2-34***

# A HYBRID DEEP LEARNING APPROACH FOR GLAUCOMA DETECTION

**Mrs. V. Divya Raj, P. Keerthy, L. Amrutha, T. Nishitha, V. Meghana**
***Department of Computer Science & Engineering, GNITS, Hyderabad, India.***
**a) divyaraj.vavilala@gnits.ac.in**
**b) pillalamarri.keerthy@gmail.com**
**c) amruthagnits@gmail.com**
**d) nish.tirumalaraju@gmail.com**
**e) meghnaa.vadla@gmail.com**

**ABSTRACT**

Glaucoma is an eye disease that causes irreversible blindness. It often occurs due to increased intraocular pressure and damages the optic nerve. Symptoms are not recognizable during early stages, making it the most challenging aspect of glaucoma. In existing technologies, including fundus imaging, require examination by an ophthalmologist. This is a very time-consuming process and might be prone to subjective human errors. These existing models also suffer from noisy images and require very high computational power.

This project uses deep learning approaches to make glaucoma detection simpler and easier by making it automated. By using multiple CNN architectures like ResNet50 and AlexNet to compare and enhance the computational accuracy. The proposed system has been implemented on both a private dataset from LV Prasad Eye Institute and public datasets from Kaggle (ORIGA, ACRIMA, DRISHTI-GS). The quality of the retinal images has been improved by various preprocessing techniques such as noise deduction, contrast enhancement, and normalization. These techniques make CNN architecture extract features more effectively. Data augmentation and hyperparameter tuning are used for training the system to enhance generalizability.

The developed model is evaluated based on the achieved accuracy to determine its effectiveness in detecting glaucoma, in which ResNet50 has achieved 95% accuracy and AlexNet has achieved 66% accuracy. It is evident that ResNet50 has achieved greater results and is more effective and efficient in detecting the required results.

## 1. INTRODUCTION

Glaucoma is a severe eye disease that can harm the optic nerve and cause irreversible blindness if not discovered and treated quickly. Regular eye exams are necessary to prevent permanent eye injury and vision loss by detecting the condition early.

Glaucoma is a kind of eye disease that mostly affects the optic nerve under high intraocular pressure conditions; it degenerates the nerve fibers, causing irreversible vision loss, and this leads to blindness when not attended to in time because it mostly does not show any symptoms within the early stages.

The most prevalent type of glaucoma is open-angle glaucoma, which occurs when the drainage angle remains open but the flow of eye fluid is obstructed. The less common and more severe type of glaucoma is when the angle closes and the pressure builds up rapidly. Diagnosis is confirmed when the patient undergoes a dilated eye exam, and onset detection is crucial. Detection is manual, subjective, and unreliable, but automated detection is computer-based, simpler to carry out, provides faster results, and is more accurate.


DOI: https://doi.org/10.5281/zenodo.16811550

Retinal images are often captured using fundus photography and provide visual representation of the back of the eye.These images are essential for diagnosing various eye diseases such as diabetic retinopathy and glaucoma [1]. To detect glaucoma, optic nerve head and the cup-to-disc ratio are important indicators, making retinal images an essential input for automated models [2].

Automation is the use of computer algorithms to perform tasks such as image processing, disease diagnosis, and detection. Evolution of fundus images periodically reduces the workload of ophthalmologists. Automated techniques facilitate mass deployment, promote efficiency, and prevent human error [3].

Before feeding machine learning algorithms, raw images undergo preprocessing methods that enhance their quality. The preprocessing methods for retinal images consist of three steps which include normalization, contrast correction and resizing, and noise reduction [4]. The CNN requires these methods to perform feature extraction effectively because it works with images acquired under diverse lighting and scenario conditions [5].

To artificially enhance the size and variability of a training set, we employ data augmentation that carries out operations like rotation, flipping, zooming, and change of brightness [6]. It enhances the resilience and generalization ability of deep learning models, making them less susceptible to real-world clinical data variability [7].

Image recognition tasks are a great fit for Convolution Neural Networks (CNNs), a type of deep learning architecture. They can automatically learn hierarchical features from input images through convolutional and pooling layers, detecting patterns like edges, textures, and complex structures [8]. CNNs are capable of learning fine distinctions in retinal images, i.e., optic nerve injury, as well as classifying with high accuracy.

## 2. RELATED WORK

Computer-aided diagnosis of glaucoma has been achieved in recent years using techniques such as the Random Forest Classifier, which is more effective than statistical techniques such as LASSO and ridge regression. It has a drawback of being dependent on certain image features and lacking higher-level patterns. Another technique, the Multi-Context Deep Network (MCDN), improves decision-making by considering whole images and high-level data but is limited to data from a particular source.

Innovations have indeed been realized, but the systems remain vulnerable to inherent limitations, especially with absolute glaucoma diagnosis automation. True, most models are trained on a single dataset, and it is difficult to train on images captured using different hospitals, imaging devices, or patients. Besides, standard practices such as Random Forest have extensive reliance upon manually designed features and are incapable of automatically discovering high-level patterns. Those are actually discoverable by novel structures like CNNs.

These systems are not scalable, cannot process large datasets or be run in real-time, and thus are not applicable to hospitals or mass screening. The project remedies this by enhancing image quality through pre-processing methods such as noise removal, contrast stretching, resizing, and normalization for improved model training. AlexNet and ResNet50 will utilize high-level feature extraction to improve detection accuracy. Performance will be measured using metrics such as accuracy, sensitivity, specificity, and F1 score. An easy-to-use platform will be developed for image uploads, real-time classification, and user interaction to provide a scalable, efficient system for automated glaucoma diagnosis.

## 3. ARCHITECTURE

The architecture of the glaucoma detection system integrates preprocessing techniques and deep learning models to deliver accurate and reliable results.

### A. Preprocessing Techniques

- **Image Resizing:** Resize all images to 256×256 pixels to make it standard and model-friendly.


DOI: https://doi.org/10.5281/zenodo.16811550

- **Data Augmentation:** Used flipping, rotation, contrast, and translation for variance augmentation and prevention from overfitting.
- **Histogram Equalization (HE):** Improved global contrast for better visibility of retinal structure.
- **CLAHE:** Preserved local contrast with preserved finer retinal details.
- **Shuffling:** Random shuffling of the data set to prevent bias and obtain a properly balanced distribution.

**B. Model Development**

For feature extraction, two deep learning models, ResNet50 and AlexNet, were used to process retinal images and extract significant patterns for glaucoma classification.

- **AlexNet:** AlexNet is a deep learning network which is employed for classifying images through the utilization of neural networks. AlexNet employs convolutional layers to learn the patterns from the images automatically and is recognized for being easy but efficient in classifying the images correctly. It demonstrated an accuracy level of 66%.
- **ResNet50:** ResNet50 is a 50-layer deep convolutional neural network that utilizes skip connections to counter the vanishing gradient problem. These residual connections enable the network to train deeper models without experiencing a decline in performance. It performed best with an accuracy rate of 95%.

## 4. IMPLEMENTATION

### 4.1 Dataset

The dataset consists of 1,335 retinal fundus images, classified in three types:

- Mild glaucoma - 131 images
- Moderate glaucoma - 887 images
- Severe glaucoma - 317 images

These sets of data contain a varied set of images that consist of presence and various stages of glaucoma.

### 4.2 Technologies Used

- Python simplifies deep learning, image processing, and numeric computations using libraries such as TensorFlow, OpenCV, and NumPy.
- TensorFlow and Keras are used for optimizing models, OpenCV for image processing operations, and NumPy for matrix operations.
- Metrics like accuracy, sensitivity, specificity, and F1 score are employed for overall assessment of model performance.
- Jupyter Notebook enables interactive coding, VS Code handles scripts, and Streamlit provides an interface for model interaction that is user-friendly.

### 4.3 Algorithm - AlexNet

AlexNet, developed by Alex Krizhevsky with Ilya Sutskever and Geoffrey Hinton, has 8 layers—5 convolutional and 3 fully connected. It processes images sized 224x224x3 (height x width x RGB). Initially, images were resized to 227x227, but the difference is negligible.

The layers 1-5 use filters (kernels) to extract features such as edges, corners, and textures. The initial layer employs 96 11x11-sized filters with a stride of 4. Subsequent layers employ smaller filters (5x5 and 3x3) and more filters in order to pick up more sophisticated patterns. They are tasked with auto-learning of features to classify.

Following every convolutional operation, ReLU (Rectified Linear Unit) is used. ReLU adds non-linearity to the model, allowing it to learn more sophisticated patterns. It's also computationally cheap and prevents the vanishing gradient problem. In the previous layers, LRN (local response normalization) was employed to simulate the lateral inhibition present in biological neurons.

The final three layers (layers 6 to 8) are fully connected; that means each neuron is connected to all neurons of the previous layer. These layers do the actual classification.

### 4.4. Algorithm - ResNet50

ResNet50 is used to extract features from the retinal image efficiently even though it is plagued with issues such as vanishing gradient problems. This is achieved by the use of residual connections (skip connections) that enable the network to capture important information and enhance training stability.

The architecture begins with a starting convolutional block, in which low-level features of edges and textures are extracted from the retinal image by a 7×7 convolution layer. Batch normalization for stabilizing training and a max pooling layer for decreasing spatial dimensions of feature maps without losing valuable information follow this.

ResNet50 is made up of four stages of residual blocks, each with bottleneck residual units having three convolution layers: 1×1 to reduce dimensions, 3×3 to extract features, and 1×1 again to restore dimensions. Skip connections help vital retinal aspects such as optic nerve head contour, cup-to-disc ratio, and nerve fiber thickness get maintained. With images progressing further into the network, more sophisticated glaucoma severity patterns are learnt effectively.

## 5. RESULTS

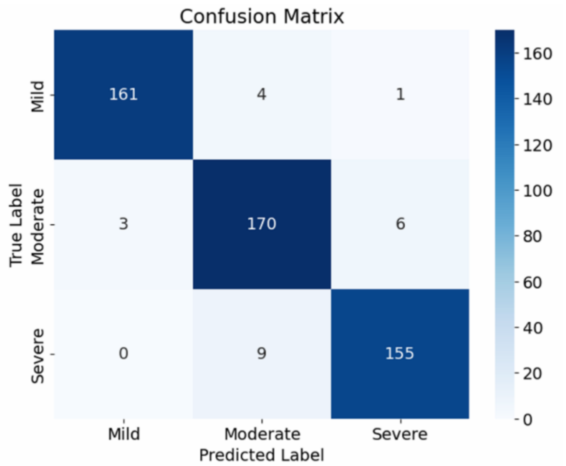
The glaucoma diagnosis framework employed AlexNet and ResNet to categorize fundus images with mild, moderate, and severe glaucoma. ResNet performed better in accuracy, feature extraction, recall, precision, and mean F1-scores compared to AlexNet, which performed poorly on medical data and often misclassified.

### 5.1.1 ResNet50 Results

The model had 95% accuracy on test data, which exhibited robust classification of all the classes. Precision was excellent with good identification of mild, moderate, and severe cases, producing few false-positive errors. Recall values remained stable, reflecting good identification of actual cases at different levels of severity. F1-scores also validated the consistency and good performance of the model, which showed good training of the dataset.

| Class | Precision | Recall | F1-Score | Support |
|---|---|---|---|---|
| Mild | 0.98 | 0.97 | 0.98 | 166 |
| Moderate | 0.93 | 0.95 | 0.94 | 179 |
| Severe | 0.96 | 0.95 | 0.95 | 164 |
| Accuracy | | | 0.95 | 509 |
| Macro Avg | 0.96 | 0.95 | 0.96 | 509 |
| Weighted Avg | 0.96 | 0.95 | 0.95 | 509 |

The confusion matrix illustrates that the model correctly classified glaucoma severity in the database, identifying 161 mild cases, 170 moderate cases, and 155 severe cases.




### 5.1.2 AlexNet Results

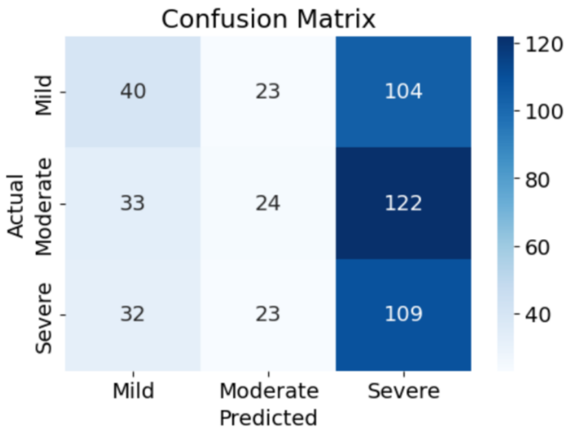
The AlexNet model was found to be not very effective, with 61.76% test accuracy as expected from train and validation sets. Precision varied from 38% for mild to 34% for moderate and 33% for severe, which suggests high false positive rates. Recall was low at 24% for mild, 13% for moderate, and 66% for severe, which shows poor classification of mild and moderate cases. F1-scores were 29% for mild, 19% for moderate, and 44% for severe, with a weighted average of 30%, highlighting substantial misclassification problems at all levels of severity.

| Class | Precision | Recall | F1-Score | Support |
|---|---|---|---|---|
| Mild | 0.38 | 0.24 | 0.29 | 167 |
| Moderate | 0.34 | 0.13 | 0.19 | 179 |
| Severe | 0.33 | 0.66 | 0.44 | 164 |
| Accuracy | | | 0.34 | 510 |
| Macro Avg | 0.35 | 0.35 | 0.31 | 510 |
| Weighted Avg | 0.35 | 0.34 | 0.30 | 510 |

The confusion matrix states that the model correctly classified glaucoma severity in the database, classifying 40 cases correctly as mild, 24 as moderate, and 109 as severe.

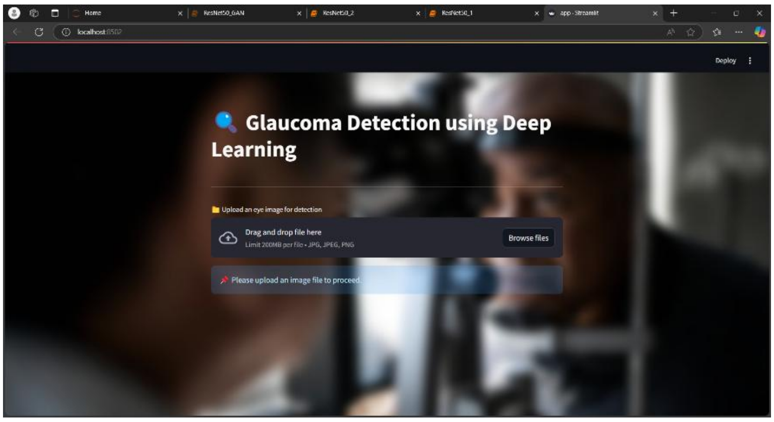

DOI: *https://doi.org/10.5281/zenodo.16811550*

*Volume-1, Issue-2, July 2025* ISSN:**3107-4308** *Paper ID:* ***ETDT-V1-I2-34***

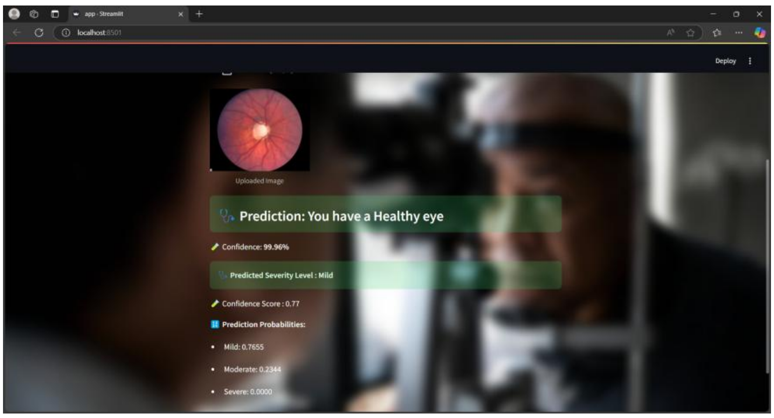




### 5.1.3 User Interface

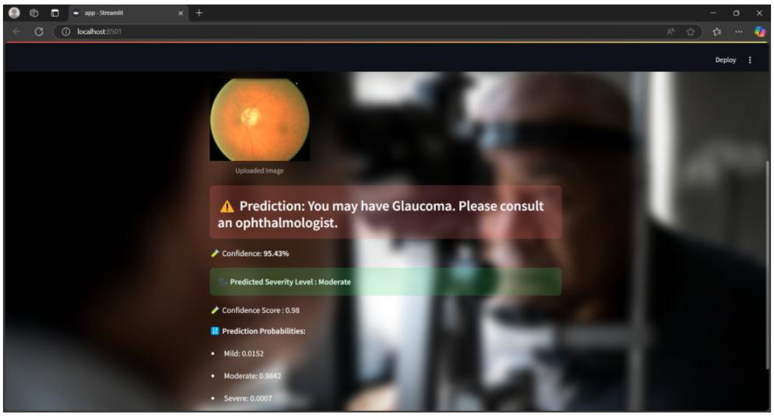
The User interface for Glaucoma Detection based on Deep Learning utilizes ResNet as the base model since it performs better than AlexNet in image classification. It offers users the convenience of uploading retinal images in JPG, JPEG, and PNG formats to predict glaucoma severity level—mild, moderate, or severe. Confidence scores for every category and an overall score ensure accurate and trustworthy predictions.



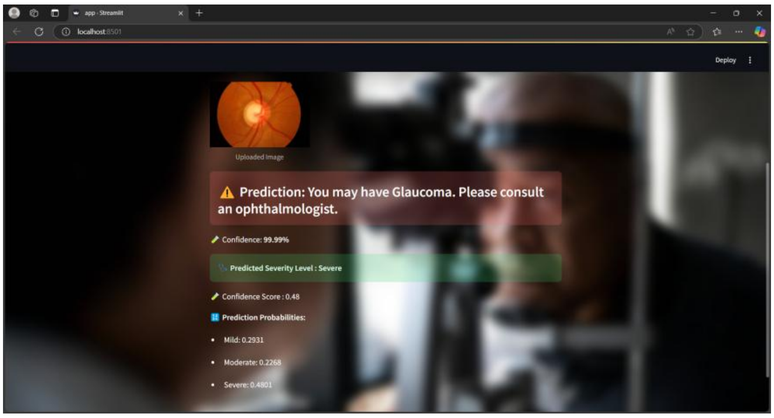

The results indicate that the eye is healthy with a severity level categorized as mild. Additionally, a confidence score of 0.92 for predicting as healthy, a confidence score of 0.77 for predicting as mild, and class probabilities are displayed.



The detected test results indicate a moderate stage of glaucoma has developed in the eye. The results display a glaucoma prediction score of 0.95 in combination with a moderate 0.98 classification score and class probabilities.




The test results demonstrate that the eye displays symptoms of severe glaucoma. The results demonstrate a 0.99 probability for glaucoma prediction and a 0.48 probability for severe prediction with class probabilities.




## 6. CONCLUSION

The ResNet model achieved the highest performance at 95% accuracy while delivering better feature extraction and maintaining stable classification performance. The severity level detection ability of ResNet makes it highly suitable for solving other complex medical imaging problems. AlexNet demonstrated poor performance through its 61.76% accuracy while displaying frequent misclassification issues that mainly affected classes mild and moderate.

The study demonstrates how deep learning-based systems function when used for ophthalmology because they offer reliable solutions for early glaucoma detection and classification. The integration of precise ResNet models in the detection process enables healthcare providers to identify urgent cases while reducing diagnostic time delays. The implementation of such systems improves glaucoma care delivery which results in better patient outcomes.

## 7. REFERENCES

**[1]** D. D. A. Parikh, R. S. Nair, "A review on retinal image processing techniques for glaucoma diagnosis," *Biomedical Signal Processing and Control*, vol. 68, 2021.

**[2]** M. S. Abramoff, M. K. Garvin, M. Sonka, "Retinal imaging and image analysis," *IEEE Reviews in Biomedical Engineering*, vol. 3, pp. 169–208, 2010.

**[3]** S. S. Li, Y. Wang, et al., "Automatic glaucoma detection using deep learning techniques," *IEEE Access*, vol. 9, pp. 45492–45504, 2021.

**[4]** B. Chen, J. Liu, "Preprocessing methods for improving deep learning performance in medical imaging," *Healthcare Technology Letters*, vol. 6, no. 2, pp. 45–50, 2019.

**[5]** H. Fu, J. Cheng, et al., "Disc-aware ensemble network for glaucoma screening from fundus image," *IEEE Transactions on Medical Imaging*, vol. 37, no. 11, pp. 2493–2501, 2018.

**[6]** Y. LeCun, L. Bottou, et al., "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.

**[7]** S. Shorten, T. M. Khoshgoftaar, "A survey on image data augmentation for deep learning," *Journal of Big Data*, vol. 6, no. 1, 2019.

**[8]** A. R. Jayanthi, R. Murugan, "Medical image data augmentation techniques: A review," *Computer Methods and Programs in Biomedicine*, vol. 195, p. 105568, 2020.