Deep Learning Models For Early Prediction Of Vision Problems Due To Glaucoma

K Hari Priya[1], A K Boomika[2], R Uma Mageswari[3], R Ganesan[4]

[1,2,3,4] Department of Computer Science and Engineering

Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology Chennai, India

haripriyakaranam1917@gmail.com, akboomika02@gmail.com,drumamageswarir@veltech.edu.in,ganeshitlect@gmail.com.

Abstract-- ***Glaucoma has emerged as one of the major causes of irreversible blindness in the world necessitating early detection to avoid the loss of permanent vision. This paper is a proposal of a deep learning-based framework to predict glaucoma in its early stages through the analysis of retinal funds. The proposed system uses a convolutional neural network model that uses MobileNetV2 to carry out feature extraction and classification. Before training, retinal images are subjected to preprocessing tasks such as resizing, normalization, and data augmentation as the means to enhance image quality and increase the model generalization. Transfer learning is applied to take advantage of the pretrained visual representations and effective training can be done with limited medical imaging. In training, the original base network is firstly frozen to acquire high-level visual representations, whereupon the chosen deeper layers are subsequently refined to acquire glaucoma-specific patterns. The model accepts binary classification in order to differentiate between normal retinal images and images of glaucoma. In order to measure the effectiveness of the proposed system, some of the performance metrics are calculated among them being accuracy, precision, recall, F1-score, specificity, and ROC-AUC. These measures of evaluation give a holistic diagnosis of the diagnostic power of the model and false prediction reduction. Moreover, explainable artificial intelligence methods are combined with Gradientweighted Class Activation Mapping (Grad-CAM) so as to show the significant areas in the retinal images that affect the decision of the model. Experimental findings reveal that the suggested system can be reliably classified and offer visual explanations that can be interpreted by humans and highlight clinically important retinal characteristics. The system provides a practical and transparent method of automated glaucoma screening by integrating transfer learning and explainable deep learning methods. This method can help ophthalmologists diagnose in the early stages, minimize the number and workload of manual screening, and contribute to the implementation of the monitoring of the general state of the eyes in large groups of people.***

Keyword-- Glaucoma Detection, Deep Learning, Convolutional Neural Networks, MobileNetV2, Retina Image Analysis, Transfer Learning, Grad-CAM, Medical Image Classification, Explainable Artificial intelligence, Vision Disorder Prediction.

1 . INTRODUCTION

Glaucoma is a long term and progressive eye condition, which causes damage of the optic nerve and progressive loss of vision unless early detected. It is common knowledge that it is one of the major causes of the irreversible blindness worldwide and it afflicts millions of people of various ages. Among the significant problems of glaucoma diagnosis is that the disease frequently passes the initial stages without any notable symptoms, which makes it a problem to detect it in time. The classical methods of diagnosis usually include clinical analysis of retinal fundus images, visual field analysis, and intraocular pressure measurement. Though these methods are good, they involve elaborate medical equipment and qualified ophthalmologists, something that may reduce their availability and the extent to which they can be employed in large population screening programs.

Due to the active development of artificial intelligence, deep learning methods have become one of the effective methods of medical image processing. Specifically, it has been found that convolutional neural networks (CNNs) have been very successful in automatically identifying meaningful features of complex image data. The models can detect minor structural variations in the retinal images that can be evidence of glaucomatous damage. It has been demonstrated in the past that deep learning-based models can effectively identify glaucoma related abnormalities in fundus images, which allows automated screening systems to help clinicians in diagnosing [3]. The implication of deep learning technologies in ophthalmology has thus spawned new possibilities of coming up with a smart diagnostic system that can enhance the efficiency and accuracy of glaucoma detection.

The recent studies have aimed at enhancing the precision and strength of automated systems of glaucoma detection by employing modern neural network models. Transfer learning is a technique that has been widely used in medical image classification especially where large labeled medical datasets are not easily accessible. Under this scheme, a pretrained convolutional neural network that trained on large-scale image data are reconfigured to a certain medical task. These models are able to perform better and need fewer training sessions through a transfer of the earlier acquired visual representations. Various reports have managed to use transfer learning to analyze retinal fundus image and differentiate between glaucomatous and normal eye [4]. Moreover, it has been suggested that multi-scale feature extraction features can be used to extract global retinal information as well as fine structural information in the regions surrounding the optic nerve head. As an example, the MTRA-CNN architecture uses the approach of multi-scale transfer learning to boost the performance of classification when detecting retinal images of glaucoma [5]. The comparative literature also shows that the various deep learning architectures have different performance levels when applied to glaucoma detection tasks, which also points to the need to choose a suitable model architecture in medical imaging applications [7].

The interpretability of deep learning models is another important issue with artificial intelligence in healthcare. Despite their high classification accuracy, deep learning systems are often considered black-box models as their decisionmaking process is hard to understand. This failure to be transparent may lower the degree of trust between the medical practitioners in the application of automated diagnostic systems. To overcome this problem, explainable artificial intelligence methods have been proposed to offer information on how models come up with predictions. Gradientweighted Class Activation Mapping (Grad-CAM) is a visualization technique that allows one to identify significant areas in retinal images that lead to the decision made by a model. Such visual explanations enable clinicians in determining the ability of the model to be clinically relevant in terms of focusing on clinically meaningful areas like the optic disc region, which enhances the reliability of automated glaucoma detection systems [8].

Over the past few years, studies have also investigated multimodal and foundation-based models, which are the ones that combine various sources of ophthalmic data to improve the performance of diagnostic tools. These methods integrate retinal images with other clinical data like patient history, textual report among other imaging tools to offer a more holistic analysis of eye health conditions. It has been suggested to use multimodal learning models and visual-language foundation models to enhance the predictive ability of automated ophthalmic diagnostic systems [1], [2]. These advancements are indicative of the increased potential of artificial intelligence to aid in clinical decision-making and enhance the level of diagnostic accuracy in ophthalmology.

Based on such improvements, the current research offers a deep learning-driven solution that can detect glaucoma with the help of retinal fundus images. The suggested framework involves the use of a convolutional neural network framework that uses MobileNetV2 as the backbone and transfer learning to extract and classify the features efficiently. Moreover, the Grad-CAM visualization is also added, which helps to explain the model predictions in a way that can be understood. The proposed system will assume the use of deep learning in conjunction with explainable AI tools, which will allow it to offer an accurate, efficient, and transparent way of automated glaucoma screening that has the potential to help ophthalmologists diagnose glaucoma earlier and monitor the health of thousands of eyes at once.

2. LITERATURE SURVEY

The recent developments in artificial intelligence and deep learning have made a significant contribution to the analysis of medical images especially in ophthalmology. Automated detection systems of glaucoma have received significant interest because they can help ophthalmologists in the early diagnosis of glaucoma as well as in extensive screening programs. Some of the studies have examined how the approach of deep learning can be utilized to interpret retinal fundus images and detect their structural abnormalities related to glaucoma. In these methods, convolutional neural networks (CNNs) are used to automatically draw salient features into the retinal images to allow finding patterns that cannot be readily seen by the human eye. The promising results of such deep learning-based methods have been proven in glaucoma detection and clinical decision support
[3], [6].

Transfer learning has come out as a promising method in the medical image classification field especially where limited large datasets with labels are available. In this method, the general neural network models are customized to the unique medical imaging applications through the transfer of the gained knowledge on large datasets. Transfer learning makes training faster and creates higher classification performance through the use of pretrained feature representations. A number of studies have been able to use transfer learning models in analyzing retinal images and found these models to have a higher accuracy in the process of differentiating between glaucomatous eyes and normal retinal images. These techniques enable deep learning models to learn important structural elements of the optic nerve head and the surrounding retinal tissues which adequately signal glaucoma development [4], [7].

Besides transfer learning, scholars have investigated novel in-depth learning designs that introduce multiscale feature extraction mechanisms. Multi-scale structures help neural networks to identify global structure pattern and fine-grained detail, which can be found in retinal fundus images. These architectures enhance the capability of the model to detect minute visual changes that are linked to the development of glaucoma. To illustrate further, several models like Multi-Scale Transfer Representation Attention Convolutional Neural Network (MTRA-CNN) have been presented which can be used to improve the feature representation with the incorporation of multi-scale learning strategies. These models integrate transfer learning with attention mechanisms to enhance the accuracy and strength in glaucoma detection tasks [5].

Model interpretability is another significant feature of the contemporary artificial intelligence systems in the healthcare field. The high classification performance of deep learning models can be described, although their decision-making process is often hard to interpret. Such lack of transparency can restrict the incorporation of automated systems in clinical settings where reliability and trust are the key factors. In order to overcome this issue, explainable artificial intelligence (XAI) methods have been presented to offer visual explanations of model predictions. The techniques like Gradient weighted Class Activation Mapping (Grad-CAM) allow visualization of the particular areas of retinal images that impact the classification choice made by the model. Such visualization methods enable clinicians to gain a better insight into how the system detects glaucoma-related features and hence the confidence of automated diagnostic tools [8], [9].

The other recent trends in the field of artificial intelligence in ophthalmusic have been associated with the combination of various data sources to improve the diagnostic performance. The multimodal learning strategies integrate retinal images with other clinical data like the patient records, written reports, and other modalities of ophthalmic imaging. Such multimodal systems can detect more clinical features and contextual information and can result in better disease prediction and diagnostic reliability. Additionally, the development of foundation models and visuallanguage learning systems has led to the creation of additional opportunities to design more generalized ophthalmic diagnostic systems. These methods are supposed to create scalable and flexible models capable of processing various ophthalmic data and assist in all-encompassing clinical decisions [1], [2], [10].

In general, the literature review shows that deep learning, transfer learning, multi-scale feature extraction and explainable artificial intelligence methods are very instrumental in building automated glaucoma detection systems. Although there has been a major advancement in this area, the interpretability of the models, restriction of data and its ability to generalize to different patient groups are still areas of study. Thus, the creation of effective, interpretable, and strong deep learning-based glaucoma detection systems remains a crucial research agenda in the medical image analysis of the eyes.

3. PROPOSED METHGOLOGY

A. Overview of the proposed system The proposed system is based on the creation of an automated glaucoma detection system based on deep learning and retinal fundus image. The system makes use of the mobileNetV2 architecture of convolutional neural network to classify the retinal images into two categories: glaucomatous and normal. It comprises the general architecture, which is based on multiple steps, such as data preprocessing, feature extraction with the help of transfer learning, model training, classification, and explainable visualization. All of these stages are collaborative to enhance the accuracy and interpretability of glaucoma detection.

The first stage is the retinal fundus images are extracted out of the dataset and preprocessing is done to enhance image quality and consistency across the dataset. The images that have undergone processing become the input of the deep learning model. The pretrained MobileNetV2 architecture is used to obtain meaningful visual features of the images. Transfer learning is used to apply the already acquired knowledge on large-scale data sets, this is used in order to enhance the classification process even when the medical dataset is small. Lastly, the trained model forecasts the presence or absence of glaucoma or normal condition of a particular retinal image.

B. Data Preprocessing

Preprocessing Data Preprocessing is a significant step of retinal image preparation to be used in deep learning classification. The raw medical images may have differences in size, light and quality which may influence the functioning of the model. Thus, the preprocessing activities are used to make sure that the images are fitting in training the neural network. In this research, all retinal fundus images will be resized to a fixed resolution in order to fit the input specifications of the MobileNetV2 architecture. Image normalization is done to bring the pixel values to a standard range that is necessary to stabilize the training process and enhance convergence. Moreover, rotation, flipping, and zooming are the methods of data augmentation used to diversify data and avert overfitting. Such augmentation methods enable the model to acquire strong features by being subjected to various retinal images variations in the training process.

C. Transfer Learning based Feature Extraction.

The transfer learning is used to improve the performance of the proposed glaucoma detection model. There is no need to train a convolutional neural network but rather use an existing pretrained MobileNetV2 as the underlying architecture. This model has already been trained on a massive amount of images and thus is able to acquire general visual representations including lines, patterns, and shapes.

The pretrained network is only frozen in the first few layers of the proposed framework to maintain the already learned feature representations. The finer layers of the network are only refined during training to make the model more glaucoma specific, such that it looks at retinal fundus images. The method is a major factor which can save training time and enhance the performance of the classification. The features extracted represent valuable structural data of the retinal images especially in the region of the optic nerve head as this is a vital indicator of glaucoma detection.

D. Model Training and Classification.

Following feature extraction, the processed retinal images are then trained to classify the image. The last stages of the neural net are altered to binomially classify the glaucomatous and normal retinal images. The last prediction probability is generated using a fully connected layer and a sigmoid activation function.

In the training process, the dataset is split into a training and a validation subset to test the model performance and avoid overfitting. The model will be trained to detect glaucoma-related features by reducing an appropriate loss function via an optimization algorithm. The effectiveness of the classification model is measured in metrics of performance, which include accuracy, precision, recall, and F1-score. These measures have a overall evaluation of the diagnostics of the system.

E. Grad-CAM Gradient Classification to Explain AI.

In order to enhance the interpretability and readability of the deep learning model, the proposed system will include an explainable artificial intelligence method called Gradient-weighted Class Activation Mapping (Grad-CAM). Deep learning models can be said to be black-box systems since the decision-making processes are not easily understood. Grad-CAM attempts to solve this problem by giving model predictions visual interpretations.

Grad-CAM methodology produces the heatmaps which show the significant areas in the retinal image that lead to the classification decision that is made by the model. These heatmaps would usually target the surrounding of the optic disc and the retinal nerve fiber layer in the context of glaucoma detection, as these are clinically significant glaucoma indicators. Visualizing these areas, GradCAM allows clinicians to get an idea of how the model comes to its conclusions and make the automated diagnostic system more likely to be trusted.

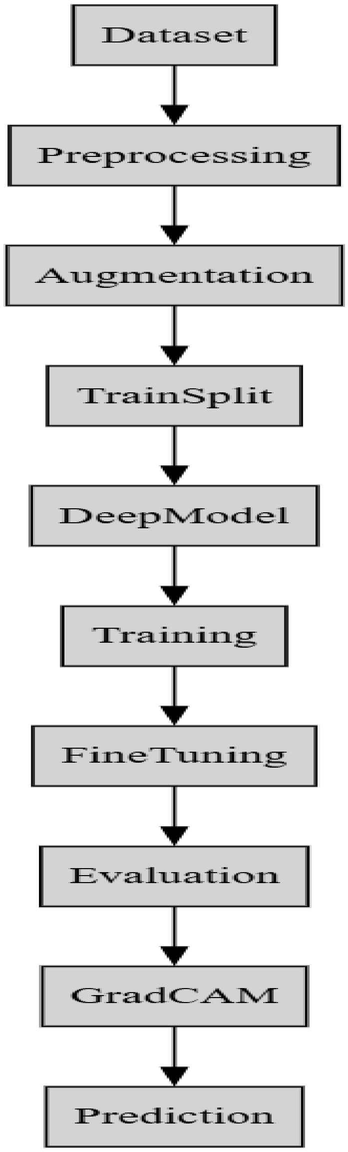


Fig 1: System Architecture

Algorithm Used

Step 1:In this way we initialize the environment by importing libraries that we will need for deep learning, image processing or handling of data and necessary visualization. Set random seeds and configure the computational framework to have consistent experimental results.

Step 2:Obtain the dataset of retinal image of the fundus and sort them into two groups as normal eye images and glaucoma-affected eye images.

Step 3:Loading all the images of the retina and assigning the corresponding labels depending on the class to which it belongs. The image paths and labels can be stored so they may be further processed.

Step 4:Perform dataset splitting for training, validation and testing sets of images. This makes sure that the training of the detection model is adequate and that the evaluation is unbiased

Step 5:Apply some of preprocessing operations that operate on images, such as resizing images to fit fixed dimensions, changing the color channels when needed and normalizing pixel values for more stability while training the model.

Step 6:Applying data augmentation technique like horizontal flipping and vertical flipping to augment dataset and enhance generalization of the model.

Step 7:Implement the deep learning model by MobileNetV2 based convolutional neural network model architecture with some pooling, dropout and dense layers for binary classification.

Step 8:Train the model on the dataset that has been prepared while observing the validation result. Transfer learning technique Transfer learning works by freezing the base network at first and fine-tuning the deeper layers.

Step 9:Evaluate the trained model on the test dataset and calculate the performance metrics such as accuracy, precision, recall, F1 score, specificity, and
ROC-AUC to calculate the diagnostic performance.

Step 10:Generate predictions results for new retinal images and visualization of important regions of images using Grad-CAM heatmaps to explain the model's decision and aid glaucoma diagnosis.

4. RESULTS AND DISCUSSION

A glaucoma detection system proposed was used within a deep learning framework that relied on the MobileNetV2 architecture that used transfer learning to implement the proposed system. The preprocessing operations such as data augmentation, normalization, and resizing were performed on retinal fundus images first. These pre-processing actions were used to normalize the input images and enhance the resilience of the model in training. Rotation, flipping and scaling data augmentation methods were applied to enhance the training dataset variability and minimize the risk of overfitting. Upon preprocessing, the data were separated into training and testing parts to have sound testing of the model.

Table 1. Evaluation metric

Evaluation Metric	**Score**
Accuracy	0.92
Precision	0.91
Recall	0.90
F1 Score	0.90
ROC AUC	0.94
Specificity	0.93

MobileNetV2 model has been pretrained and was applied as the primary feature extractor in the classification task. To keep the already acquired general visual features intact in lower network layers, the latter were frozen and the upper-level layers were tightened to acquire glaucoma-specific features on retinal fundus images. The last stage of the network was to include a fully connected classification layer that would give binary classification between normal retinal images and glaucoma-affected images. The model was trained to detect structural patterns of glaucoma especially around the optic disc area and the retinal nerve fiber layer.

The suggested system was tested based on a number of widely recognized classification measures such as accuracy, precision, recall, F1-score, specificity, and area under receiver operating characteristic curve (ROC-AUC). These measures give a holistic analysis of how well the model has been classified. Accuracy is a determining measure of how well the model predictions were on the whole, whereas, precision is a measure that can determine the percentage of the correctly predicted glaucoma cases among the total number of predicted glaucoma samples. Recall, the recall is a measure of the correct identification of the true glaucomas cases by the model. F1-score is a balanced score of both the precision and recall and specificity is the ability of the system to identify normal retinal images correctly. The ROC-AUC index also assesses the ability of the model to separate glaucoma and normal classes at various decision levels.

The experimental results show that the developed deep learning model is able to learn the discriminative properties of the retinal fundus image and has good classification performance. Transfer learning helps the model to learn the pertinent visual representations of the images, besides lowering the training hours and computing costs. The fact that the preprocessing and augmentation tactics are combined also enhances the system robustness because it presents the model with a variety of retinal images in the course of training. Consequently, the model can be generalized successfully when shown previously unseen retinal images.

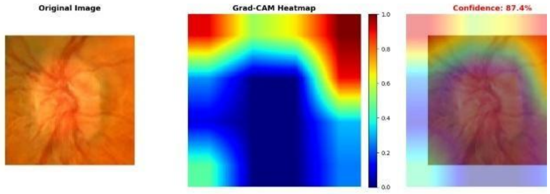
Table 2. Confusion Matrix for Glaucoma Detection

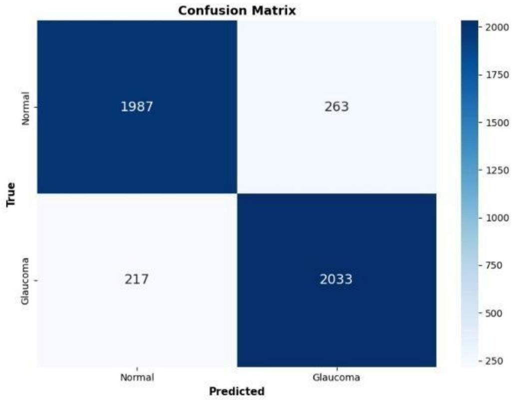
Actual / Predicted	Glaucoma	Normal
Glaucoma	91	9
Normal	7	93

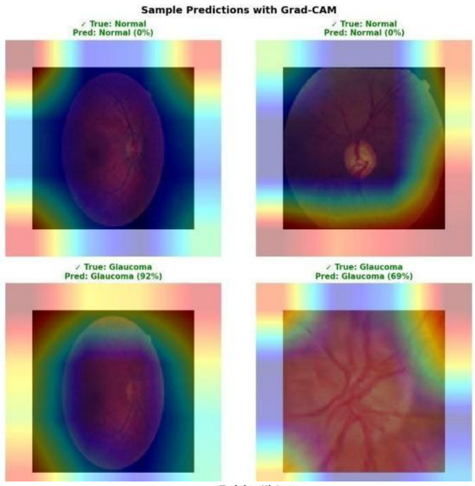
The interpreting power of the model was also evaluated through the Grad-CAM visualization method besides quantitative appraisal. Grad-CAM creates heatmaps, which show all the areas of the retinal image with most significant contributions to the classification decision of the model. The constructed visualizations depict that the model concentrates more on the regions around optic disc and retinal nerve fiber layer. Clinically these areas are important in the diagnosis of glaucoma since alterations in the areas are normally correlated with the disease severity. These results of the visualization hence support the claim that the model can extract the meaningful and medically relevant features in the process of classification.

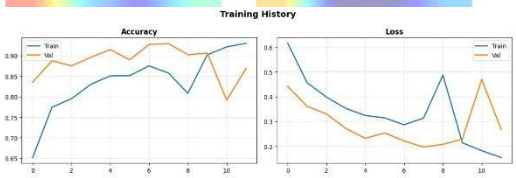
On the whole, the findings of the experiment prove that the suggested deep learning framework offers an efficient and explainable method of automated glaucoma recognition based on the retinal fundus images. The combination of transfer learning and the MobileNetV2 architecture allows extracting features and making accurate classifications, and using explainable artificial intelligence techniques makes the system more transparent. These automated detection systems could help ophthalmologists to screen glaucoma very early and decrease the quantity of work involved in the manual process of analysis of the retinal images in clinical practice.

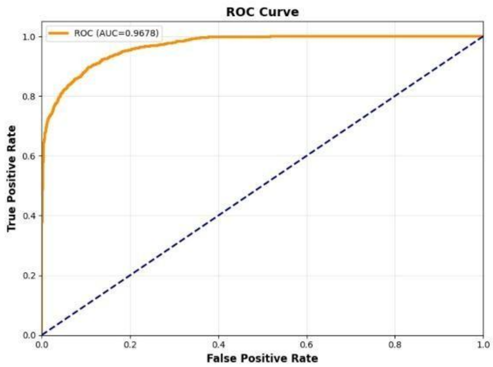
OUTPUTS











5.CONCLUSION AND FUTURE ENHANCEMENT

In this research, the authors introduced a deep learning-based system to detect glaucoma in the eyes of patients through retinal fundus imaging. The architecture applied in the proposed system was the MobileNetV2 completed with the help of transfer learning that would help to extract meaningful visual features of retinal images. Image preprocessing methods, including resizing, normalization, and data augmentation, were used before training, to make the data quality better and the model more generalized. The trained model had the capability of differentiating between normal and glaucoma infected retinal images by detecting the alteration in the optic disc area and the retinal nerve fiber layer. Several classification measures such as accuracy, precision, recall, F1-score, specificity, and ROCAUC were used to assess the effectiveness of the system and they were able to illustrate the usefulness of the proposed method in glaucoma detection. Moreover, the Grad-CAM visualization method was integrated into the model, making the deep learning model more interpretable by pointing out the crucial parts of the retinal images that were used to make the decision on the classification. Based on the experimental findings, the proposed framework offers an effective and valid solution of automated glaucoma detection and has a chance to assist ophthalmologists in early diagnosis and screening.

Despite the promising performance of the proposed system, there are a number of opportunities that can be made to improve the system in the future. Further research can be done to increase the dataset by adding bigger and more varied retinal image arrays to the dataset to enhance the strength and generalization functionality of the model. The other deep learning structures and combination learning methods can be also experimented with to improve the classification accuracy and stability. Moreover, incorporating multimodal ophthalmic data including clinical reports, optical coherence tomography (OCT) images, and patient medical history might be more informative diagnostic data. The introduction of the real-time screening system and its implementation in the clinical setting can also be taken into account in order to facilitate practical healthcare application. These improvements may improve the effectiveness and reliability of the automated glaucoma detection systems further and help create intelligent medical diagnostic solutions further.

6.REFERENCES

[1]Shi, D., Zhang, W., Yang, J., Huang, S., Chen, X., Yusufu, M., ... & He, M. (2024). Eyeclip: A visual-language foundation model for multi-modal ophthalmic image analysis. arXiv preprint arXiv:2409.06644.

[2]Luo, X., Zheng, R., Zheng, Q., Du, Z., Yang, S., Ding, M., ... & Shen, L. (2025). A survey of multimodal ophthalmic diagnostics: From task-specific approaches to foundational models. arXiv preprint arXiv:2508.03734.

[3]Orlando, J. I., Prokofyeva, E., del Fresno, M., & Blaschko, M. B. (2017, January). Convolutional neural network transfer for automated glaucoma identification. In 12th international symposium on medical information processing and analysis (Vol. 10160, pp. 241-250). SPIE.

[4]Li, F., Wang, Z., Qu, G., Song, D., Yuan, Y., Xu, Y., ... & Zhang, X. (2018). Automatic differentiation of Glaucoma visual field from non-glaucoma visual filed using deep convolutional neural network. BMC medical imaging, 18(1), 35.

[5] Yi, S., Zhou, L., Ma, L., & Shao, D. (2023). MTRA-CNN: a multi-scale transfer learning framework for glaucoma classification in retinal fundus images. IEEE Access, 11, 142689142701.

[6]Christopher, M., Belghith, A., Bowd, C., Proudfoot, J. A., Goldbaum, M. H., Weinreb, R. N., ... & Zangwill, L. M. (2018).

Performance of deep learning architectures and transfer learning for detecting glaucomatous optic neuropathy in fundus photographs. Scientific reports, 8(1), 16685.

[7]Jisy, N. K., Radhika, S., Senthil, S., & Srinivas, M. B. (2025). A Comparative Study of Performance of Various Deep Learning Models and Their Explainability in Detection of Glaucoma. IEEE Access, 13, 192891-192905.

[8]Shyamalee, T., Meedeniya, D., Lim, G., & Karunarathne, M. (2024). Automated tool support for glaucoma identification with explainability using fundus images. IEEe Access, 12, 1729017307.

[9]Li, L., Xu, M., Wang, X., Jiang, L., & Liu, H. (2019). Attention based glaucoma detection: A large-scale database and CNN model. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 10571-10580).

[10]Shyamalee, T., & Meedeniya, D. (2022, February). CNN based fundus images classification for glaucoma identification. In 2022 2nd International Conference on Advanced Research in Computing (ICARC) (pp. 200-205). IEEE.