Early Glaucoma Detection Using EfficientNet-b1 on Fundus Images

Dr. David Raj Micheal
Department of Mathematics
School of Advanced Sciences
Vellore Institute of Technology Chennai
Tamil Nadu – 600127
davidraj.micheal@vit.ac.in

Balanandha R
Department of Mathematics
School of Advanced Sciences
Vellore Institute of Technology Chennai
Tamil Nadu – 600127
balanandha.r2023@vitstudent.ac.in

***Abstract*—Early glaucoma screening is explored by processing low-quality fundus images using the EfficientNet-b1 model. The study employs anomaly detection techniques to identify glaucoma at an early stage, even in images with noise or artifacts, using a dataset of 4000 fundus images. To improve detection performance, the approach integrates image enhancement methods with EfficientNet-b1, enabling the system to adapt based on the quality of the images analyzed. The goal is to make glaucoma screening more reliable, particularly in resource-constrained settings where imaging devices may produce lower-quality outputs. By addressing issues related to image quality, the system can generalize across varying conditions, ensuring robust performance in real-world applications. This work highlights the importance of early diagnosis and the potential of using deep learning to improve outcomes, ultimately contributing to the advancement of glaucoma detection and reducing the risk of vision loss.**

***Index Terms*—Glaucoma Screening, EfficientNet-b1, Anomaly Detection, Fundus Images, Image Enhancement, Low-Quality Images, Early Detection, Deep Learning, Resource-Constrained Settings.**

I. INTRODUCTION

Glaucoma, a leading cause of irreversible blindness, is often diagnosed at advanced stages due to its asymptomatic nature. Early detection is crucial, particularly in resource-constrained settings where imaging quality can vary significantly. This research proposes a novel approach for early glaucoma screening by leveraging anomaly detection with deep learning models, specifically designed to handle low-quality fundus images. Using a dataset of 4000 fundus images, the study combines image enhancement techniques with adaptive Convolutional Neural Networks (CNNs) to improve detection accuracy despite noise and artifacts. By enabling the system to learn from the state of images, the proposed method enhances the reliability of glaucoma detection, thus addressing a key challenge in practical, real-world settings where diagnostic tools must adapt to varied imaging conditions. This advancement in anomaly detection and adaptive deep learning holds the potential to significantly improve early glaucoma screening, making it accessible in settings with limited resources and improving patient outcomes.

II. OBJECTIVE

The primary objective of this research is to develop an advanced and adaptive deep learning system for early glaucoma screening using low-quality fundus images, leveraging the EfficientNet-b1 model to enhance both the reliability and accuracy of glaucoma detection. This system aims to improve detection performance by preprocessing and analyzing images with noise and artifacts, while adaptively handling variations in image quality to ensure robust performance across diverse real-world conditions. By utilizing EfficientNet-b1, the research seeks to increase the practicality of glaucoma screening in resource-constrained settings, enabling accurate early detection with limited imaging resources and ensuring that the model generalizes well across different image qualities for consistent diagnosis.

III. LITERATURE REVIEW

It deals with the development of a CAD system for early glaucoma detection, essentially carrying out image processing along with supervised learning. As a matter of fact, this model has outperformed recent deep models such as ResNet-50, VGGNet-16, and GoogLeNet in terms of accuracy, sensitivity, specificity, precision, and F1 score. Three ConvNet architectures are considered: GoogLeNet, VGGNet-16, and ResNet-50; transfer learning from the ImageNet dataset is used for deep feature extraction that will support distinguishing between normal and abnormal fundus images. Therefore, the proposed model performs extremely well with a high accuracy and sensitivity for both the public and private datasets of glaucoma images. [1]

Aims to design an automatic glaucoma classification technique based on deep learning models. A total of 634 colour fundus images were collected and labelled by the ophthalmologists who used the EfficientNet, MobileNet, DenseNet, and GoogLeNet models. The proposed EfficientNet-b3 model performed best, with a test accuracy of 0.9652 and F1-score of 0.9512. This was done through blood vessel segmentation of the retinal fundus images, using a U-Net model. For the model MobileNet v3, a test accuracy of 0.8348 and F1-score of 0.7957 have been achieved. [2]

The research work deals with residual networks for early detection of glaucoma by using color fundus images. The authors introduce a proprietary dataset of early-stage glaucoma fundus color images and use a pre-trained ResNet50 network trained on ImageNet. The results are that the model has a validation accuracy of 96.95 percentage, therefore deep learning algorithms might be a very inexpensive screening test for early detection of glaucoma. [3]

A GAN-based anomaly detection model was developed for ocular disease detection with fundus images. It performed quite satisfactorily with an AUC of 0.896 and a sensitivity of 82.69 percent on internal testing with 90,499 images. It yielded an AUC of 0.900, with a sensitivity of 83.25 percent, when externally tested. It had successfully detected six common diseases and various degrees of diabetic retinopathy. The results indicated that this is a promising, low-cost tool for ocular disease screening; however, more clinical validation is required.. [4]

An automated framework to detect and segment diabetic eye diseases through a Fast Region-based Convolutional Neural Network along with fuzzy k-means clustering. It consists of two steps: disease localization and segmentation. In the experimentation on benchmarked datasets, it means the performance in Intersection-over-Union is 0.95 and the mean Average Precision is higher than 0.94 for diabetic retinopathy, diabetic macular edema, and glaucoma. This method outperforms other techniques that have been so far proposed. [5]

The application of AI to the detection and monitoring of glaucoma. AI techniques, especially deep learning and machine learning, can analyze common structural and functional tests such as OCT, fundus photography, and visual field testing with sensitivity and specificity of more than 0.90 with an AROC of more than 0.90. A combination of these inputs will enhance the diagnostic accuracy, and AI may detect the progression of glaucoma more sensitively than current clinical methods can. While the fundus photographs present a basic screening modality, in cases with high risk, it is recommended to perform more sophisticated investigation techniques. AI will be able to support clinical decision-making and efficiency of care; however, there are challenges ahead, such as explainable AI and also regulatory approval. [6]

An InceptionResNetV2 deep learning system was trained for the diagnosis of GON in ultra-widefield fundus images. This system tested with images from 10,590 subjects reached AUC values ranging between 0.983 and 0.999, with sensitivities of 97.5-98.2 percent, and specificities of 94.3-98.4 percent, performing comparably to experienced ophthalmologists, thus being able to increase glaucoma screening and the possibility of early diagnosis. [7]

A new approach in the design of an automated glaucoma detection system using fundus images with an architecture based on a support vector machine and comprising the VGG-19 network. Glaucoma is a neurodegenerative disorder causing loss of vision, generally diagnosed by measuring the intraocular pressure and Cup-Disc-Ratio. Thus, the proposed system classifies the images based on a threshold of CDR at 0.41, wherein CDR values greater than 0.41 will indicate an image as glaucoma-affected and vice-versa. This methodology was tested on 175 digital color fundus images, where it achieved a precision of 94 percent in classification. [8]

The application of DenseNet combined with DarkNet for the classification of normal and glaucoma-affected fundus images. Glaucoma is responsible for the maximum number of blindness scenarios due to damage in the optic nerve head area. Due to the asymptomatic nature of the disease, conventional techniques like tonometry, ophthalmoscopy, and perimetry are not effective to diagnose it. While Densenet-DarkNet achieved an accuracy of 99.7 percent, sensitivity of 98.9 percent, and specificity of 100 percent on the HRF dataset; 89.3 percent accuracy, with 93.3 percent sensitivity and 88.46 percent specificity on RIM-1; and 99, 100, and 99 percentage ACRIMA, respectively. [9]

A computer-aided diagnosis system that adopts automated early eye disease detection from retinal fundus images. It involves a hybrid ensemble approach: iAlexNet features coupled with ReliefF for dimensionality reduction and XgBoost as the classification algorithm. Class balancing is realized by augmentation techniques. This deep-ocular model, therefore, proposed based on the AlexNet-ReliefF-XgBoost framework, achieved an accuracy of up to 95.13 percent, hence proving that this framework has immense potential for early diagnosis and screening of ocular diseases. [10]

The deep learning model, Faster R-CNN, to detect glaucoma in fundus images. The presented study applies the transfer learning with architectures, that is, ResNet50 and VGG16, where it illustrates up to 96 percent accuracy with ResNet50. The method is tested on two datasets, DRISHTIGS and ORIGA, adding up to a total of 751 images. This indicates that it could be used for large-scale glaucoma screening programs. The performance for ResNet50 is the best among the others. [11]

The proposed study is based on a fine-tuned ResNet-50 architecture on more than 5,000 retinal images from the ODIR dataset for the detection and classification of eye disorders such as glaucoma, cataract, and diabetic retinopathy. Overfitting has been prevented by several data augmentation techniques to enhance generalization. The performance of the models has been impressive with accuracy at 92.60 percent, precision at 93.54 percent, recall at 91.60 percent, an F1-score at 91.68 percent, which goes to prove that deep learning can perform early diagnosis automation and improve patient outcomes. [12]

It was a CNN-based early glaucoma detection from fundus images. It used the base models of AlexNet, VGG16, ResNet50, and InceptionV3, further combined into ResNet50 and InceptionV3 to come up with a hybrid model that could work better. When tested on the ORIGA, STARE, and REFUGE datasets, the model achieved high F1 Scores amounting to 97.4, 99.1, and 99.2 percentage, respectively. The system introduced ensures a reliable tool, which is more accurate in the diagnosis of glaucoma and, through early

detection, enhancing the efficiency of operations of medical screenings. [13]
It not only performs a comparative analysis of the performances of ResNet101 and MobileNet, but also a refined VGGNet-19 for glaucoma diagnosis on fundus images. After evaluating the overall performance on a test set, the ensemble approach gave an overall performance of 94 percent, while VGG19 slightly outperforms the three classes in terms of F1-score. While the models were performing well, the study points out some issues that in different datasets, there are inconsistencies and there might be a potential bias; thus, more research is called for to come up with a better and more reliable model to be put into practice. [14]
Deep learning models were used for glaucoma detection and mask identification for COVID-19. For glaucoma, the Convolutional Neural Network (CNN) with max pooling achieved the best performance with 87.99 percent training and 89.11 percent validation accuracy, while the Xception transfer learning model excelled with 97.63 percent training and 98.11 percent validation accuracy. For mask detection, a transfer learning model outperformed two deep learning models. Future work will expand datasets and evaluate additional models for improved performance and comparison. [15]

IV. METHODOLOGY

A. Data Collection and Preprocessing

For this project, the dataset used is the glaucoma dataset, which is a balanced subset of standardized fundus images from the Rotterdam EyePACS AIROGS [1] set. The dataset is split into three main subsets: training (4000 images, approximately 84%), validation (385 images, approximately 8%), and test (385 images, approximately 8%). The training set contains two distinct classes: referable glaucoma (RG) and non-referable glaucoma (NRG), with separate folders for each class.

The data collection phase involves obtaining these images, which are then preprocessed to ensure uniformity and suitability for model training. First, the images are resized to a consistent dimension (either 128x128 or 300x300 pixels), ensuring they meet the input requirements of the model. To improve training stability and model performance, the images are normalized, scaling pixel values to a range between 0 and 1. Furthermore, data augmentation techniques such as rotation, flipping, and zooming are applied to increase the diversity of the training data. This helps prevent overfitting and improves the model's ability to generalize. This preprocessing pipeline ensures that the dataset is clean, standardized, and augmented, providing the model with high-quality, varied input for effective learning.

B. Train Autoencoder Model

An autoencoder is a type of neural network designed to learn efficient representations of data, often used for reducing dimensions or extracting key features. It works by having two parts: an encoder and a decoder. The encoder compresses the input images into a smaller, simpler form, capturing the most important features of the image, while the decoder tries to recreate the original image from this compressed version. In this project, we use an autoencoder to learn how to efficiently represent fundus images by minimizing the difference between the original image and the one the autoencoder reconstructs. To train the autoencoder, we pass the original fundus images through the encoder, then through the decoder to reconstruct them. The training process involves adjusting the model's parameters so that the reconstructed image is as close as possible to the original, using a loss function like mean squared error (MSE) to measure how well the reconstruction matches the input. This is done using an optimizer like Adam, which helps the model gradually improve its ability to compress and reconstruct images accurately. Over time, the autoencoder learns to extract meaningful features from the fundus images, which can then be used for further tasks, like detecting glaucoma.

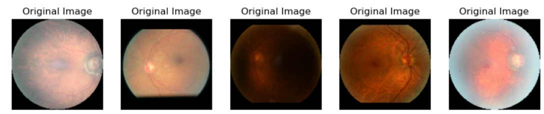

Fig. 1. Data Preprocessed

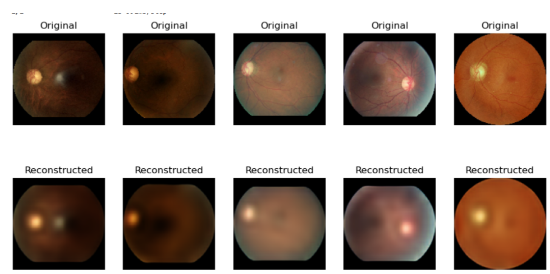

Fig. 2. Autoencoder

C. EfficientNet B1 Model

In this project, we first resize and preprocess the fundus images to ensure they are in a suitable format for the EfficientNetB1 model. Since EfficientNet requires input images of a specific size (in this case, 300x300 pixels), we resize the reconstructed images from the autoencoder to match this dimension. Additionally, we normalize the pixel values of the images to a range of 0 to 1 to ensure that the model processes the data more efficiently. This preprocessing step ensures that the input images are of consistent quality and format, improving the model's performance.

Once the images are ready, we proceed with designing the EfficientNetB1 model. EfficientNet is a highly efficient convolutional neural network architecture that has shown exceptional performance in image classification tasks, owing to its ability to scale depth, width, and resolution of the network in a balanced manner. In this project, we use the EfficientNetB1 model, which is pre-trained on ImageNet, and modify it by removing the top classification layer to fine-tune it for our specific task. We add a custom classifier on top of the pre-trained model to handle the binary classification task (i.e., classifying images into referable glaucoma and non-referable glaucoma). The model is then compiled with an appropriate optimizer, like Adam, and a binary cross-entropy loss function, which is commonly used for binary classification tasks.

Finally, we train the EfficientNetB1 model using the pre-processed training data, with the goal of teaching it to classify the fundus images into the two categories: referable glaucoma (RG) and non-referable glaucoma (NRG). During training, the model adjusts its weights based on the loss function, gradually improving its ability to accurately predict the class of each input image. We use the training and validation data to monitor the model's performance and ensure it generalizes well to unseen data, while also preventing overfitting by employing techniques like early stopping. The training process helps the model learn from the features extracted by the autoencoder, leading to better predictions for the glaucoma detection task.

V. Evaluation

This section describes the proposed model's experimentation setup and evaluation procedure.

A. Experimentation Setup

The methods proposed in this research are initially trained on the provided dataset. The models (GRU and LSTM) are trained using the processed features from the UCI HAR dataset, with the training and test sets already separated. The models undergo several epochs of training, and the evaluation is performed after each epoch using the validation data.

B. Evaluation Measures

When assessing models for multi-class classification, various important metrics are considered to thoroughly gauge their performance. In addition to standard metrics like accuracy and loss, precision, recall, F1 score, and confusion matrices are also analyzed to better understand the model's classification abilities.

1) Accuracy: Accuracy indicates the percentage of correctly classified samples out of the total, providing a general idea of the model's predictive accuracy. It is calculated using the following formula:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \tag{1}$$

where:

- TP = True Positives
- TN = True Negatives
- FP = False Positives
- FN = False Negatives

2) Loss: Loss measures the discrepancy between predicted and actual labels, indicating model performance during training. Lower values suggest better performance. Loss is computed using categorical cross-entropy, with lower values signifying a better-performing model.

3) Precision: Precision represents the ratio of correctly predicted positive cases to all cases predicted as positive, indicating how well the model avoids false positive predictions. It is calculated using the following formula:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{2}$$

4) Recall: Recall, also known as sensitivity or the true positive rate, is the percentage of positively confirmed true cases among all positive test results. It indicates the model's effectiveness in capturing positive cases. The formula for Recall is:

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{3}$$

5) F1-Score: The F1 score is the harmonic mean of Precision and Recall, providing a balanced view of model performance, especially useful for imbalanced datasets. The F1 score combines both false positives and false negatives into a single metric, calculated as:

$$\text{F1 Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{4}$$

VI. Results and Analysis

The graph shows the training and validation loss of the autoencoder model over 5 epochs, reflecting how well the model has learned to reconstruct images with minimal error. Initially, both the training loss (blue line) and validation loss (orange line) start relatively high, with a sharp decrease after the first epoch, followed by a more gradual reduction. By the last epoch, the model achieves a very low training loss of approximately 0.00085772 and a validation loss of 0.00071728, indicating effective reconstruction with minimal error. The close alignment of training and validation loss throughout the epochs suggests the model generalizes well, as there is no sign of overfitting. The final low values for both metrics indicate that the autoencoder has successfully learned the essential features needed to represent the original images accurately.

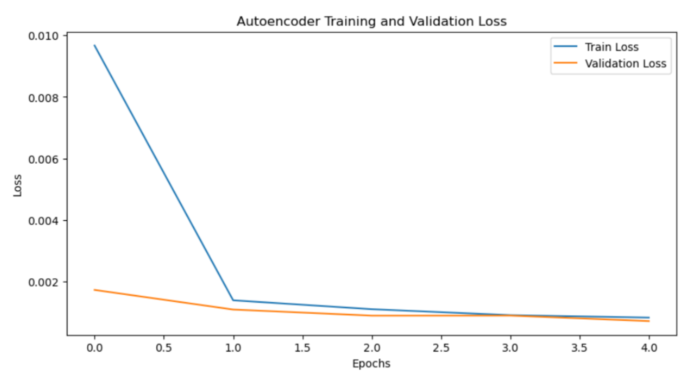

Fig. 3. Accuracy and loss of Autoencoders

The confusion matrix illustrates the performance of the EfficientNet model in classifying glaucoma fundus images into two categories: referable glaucoma (RG) and non-referable glaucoma (NRG). In this matrix, the model correctly identified 220 images as NRG and 165 images as RG, representing the true positives and true negatives for each category. The overall accuracy achieved by the model is 86.22%, with a training loss of 0.3363, indicating a good level of performance. Additionally, the model shows strong validation metrics, with a validation accuracy of 88.28% and a lower validation loss of 0.2628, suggesting it generalizes well to unseen data. The balance between training and validation metrics indicates that the model effectively learned features relevant for distinguishing between RG and NRG images, while the relatively low validation loss further confirms the model's robustness. This matrix helps visualize the instances where the model succeeded or failed in classification, guiding future adjustments for improved accuracy.

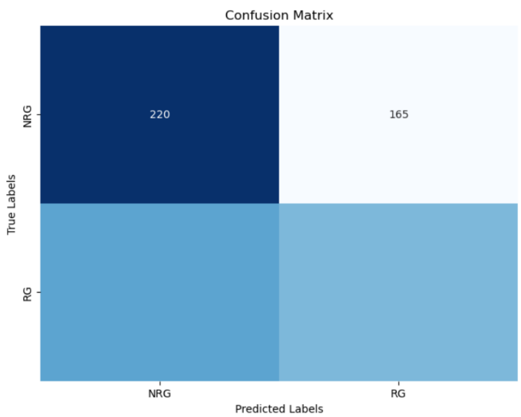

Fig. 4. Confusion Matrix of EfficientNet B1

The EfficientNet-B1 model achieved a training accuracy of 86.22% with a training loss of 0.3363, and a validation accuracy of 88.28% with a validation loss of 0.2628, indicating a well-generalized performance in distinguishing between referable glaucoma (RG) and non-referable glaucoma (NRG) cases. The slightly higher validation accuracy compared to training accuracy suggests minimal overfitting, while the low loss values indicate effective learning of relevant features. Further fine-tuning, such as increasing the training epochs, adjusting the learning rate, and employing additional data augmentation, could enhance model performance. The use of precision, recall, and F1-score analysis, along with a confusion matrix, would provide a deeper understanding of the model's reliability in correctly classifying RG and NRG cases, ensuring robust clinical applicability for glaucoma detection.

VII. Conclusion

In conclusion, this research demonstrates the effective application of an EfficientNet-B0 model enhanced by an autoencoder for accurate glaucoma detection from fundus images. The autoencoder efficiently learned to reconstruct images with minimal error, providing essential features for the EfficientNet classifier. With a training accuracy of 86.22% and a validation accuracy of 88.28%, along with low training and

Fig. 5. Prediction of EfficientNet B1

validation loss values, the EfficientNet-B0 model exhibited strong generalization and reliability in distinguishing between referable glaucoma (RG) and non-referable glaucoma (NRG) cases. The confusion matrix further highlighted the model's classification accuracy, suggesting minimal overfitting and robust feature learning. The results indicate that this model can be a valuable tool in clinical settings for early glaucoma detection, contributing to timely diagnosis and intervention. Further improvements, including precision, recall, and F1-score analysis, and adjustments to hyperparameters or data augmentation techniques, could enhance the model's robustness and reliability, paving the way for real-world deployment in automated glaucoma screening systems.

References

[1] Joshi, Partibane, et al. Glaucoma Detection Using Image Processing and Supervised Learning for Classification. Journal of Healthcare Engineering, 2022.

[2] Islam, Mashfu, et al. Deep Learning-Based Glaucoma Detection With Cropped Optic Cup and Disc and Blood Vessel Segmentation. IEEE Access, 10, 2828-2841,2022.

[3] Ovreiu, Cristescu, et al. Early Detection of Glaucoma Using Residual Networks. 2020 13th International Conference on Communications (COMM), 161-164,(2020).

[4] Han, Li, et al. Application of an Anomaly Detection Model to Screen for Ocular Diseases Using Color Retinal Fundus Images: Design and Evaluation Study. Journal of Medical Internet Research, 23,(2021).

[5] Nazir, Irtaza, et al .Retinal Image Analysis for Diabetes-Based Eye Disease Detection Using Deep Learning. Applied Sciences,(2020).

[6] Mursch-Edlmayr, Ng, et al. Artificial Intelligence Algorithms to Diagnose Glaucoma and Detect Glaucoma Progression: Translation to Clinical Practice. Translational Vision Science and Technology,(2020).

[7] Li, Lin, et al, Deep learning for automated glaucomatous optic neuropathy detection from ultra-widefield fundus images. British Journal of Ophthalmology, 105, 1548 - 1554,(2020).

[8] Raja, Shanmugam, et al. An Automated Early Detection of Glaucoma using Support Vector Machine Based Visual Geometry Group 19 (VGG-19) Convolutional Neural Network. Wireless Personal Communications, 118, 523 - 534, (2021).

[9] Akbar, Hassan, et al. Detection of microscopic glaucoma through fundus images using deep transfer learning approach. Microscopy Research and Technique, 85, 2259 - 2276,(2022).

[10] Abbas, Albathan, et al. Deep-Ocular: Improved Transfer Learning Architecture Using Self-Attention and Dense Layers for Recognition of Ocular Diseases. Diagnostics, 13,(2023).

[11] Ajitha,Judy . Faster R-CNN classification for the recognition of glaucoma. Journal of Physics: Conference Series, 1706,(2020).

[12] S, S.,B. K., A. Optimized Screening of Glaucoma using Fundus Images and Deep Learning. International Journal on Recent and Innovation Trends in Computing and Communication, (2023).

[13] Govindan, Dhakshnamurthy, et al. A Framework for Early Detection of Glaucoma in Retinal Fundus Images Using Deep Learning,(2024).

[14] Vijayalakshmi, Basak, et al. Glaucoma Detection through Deep Learning: A Transfer Learning Techniques using CDR Feature Extraction. 2024 Third International Conference on Distributed Computing and Electrical Circuits and Electronics,(2024).

[15] Shadin, Sanjana,et al. Performance Analysis of Glaucoma Detection Using Deep Learning Models. 2022 International Conference on Innovations in Science, Engineering and Technology (ICISET), (2022)