Volume 11, Issue 4, April – 2026 International Journal of Innovative Science and Research Technology
ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/26apr1398

Multi-Disease Detection in Retinal Fundus Images

Vijaya Jyothi Chiluka[1]; P. Jasmitha[2]; R. Srikavya[3]; S. Meghana[4]; V. Sruthi[5]

[1]Assistant Professor, CSE, [2,3,4,5]B. Tech 4th Year Student, CSE, Vignan's Institute of Management and Technology for Women, Hyderabad, India.

Publication Date: 2026/05/08

Abstract: Retinal conditions like Diabetic Retinopathy (DR), Glaucoma, and Neovascularization (NV) are some of the common reasons for permanent blindness across the globe, but their early detection is challenging due to the need for manual evaluation of retinal images, which is time-intensive and subjective and not feasible for population-level screening. This study proposes an automated deep learning system capable of classifying retinal fundus images based on the following classes: Diabetic Retinopathy, Glaucoma, Neovascularization, and Normal. The performance of three transfer learning models (EfficientNet, ResNet50, and MobileNetV2) was assessed using the APTOS 2019 dataset, consisting of around 3,500 retinal fundus images. Preprocessing methods such as resizing to 224 × 224, pixel normalization, Gaussian noise removal, and augmentation were applied to improve the performance of the proposed deep learning system. Grad-CAM was also incorporated to provide clinicians with heatmaps indicating the areas of the retina that play a crucial role in the classification task. Out of the tested architectures, EfficientNet performed better than its competitors, producing an overall accuracy of 67%, with weighted precision, recall, and F1-scores within the range of 66-67%. However, the model used to detect Glaucoma was the most converged, with a ROC-AUC score of 0.75, whereas the diabetic retinopathy model produced multi-class AUC scores ranging from 0.83 to 0.96, depending on the disease category. This study has set an important benchmark for simultaneously screening patients for multiple diseases. The main finding of this research is that dataset size and class balance are critical determinants of classification success.

Keywords: *Retinal Disease Classification, Multi-Disease Detection, Diabetic Retinopathy, Glaucoma, Neovascularization, Deep Learning, Transfer Learning, EfficientNet, ResNet50, MobileNetV2, Convolutional Neural Networks, Grad-CAM, Fundus Image Analysis, APTOS 2019, Medical Image Classification.*

How to Cite: Vijaya Jyothi Chiluka; P. Jasmitha; R. Srikavya; S. Meghana; V. Sruthi (2026) Multi-Disease Detection in Retinal Fundus Images. *International Journal of Innovative Science and Research Technology*, 11(4), 3612-3623. https://doi.org/10.38124/ijisrt/26apr1398

I. INTRODUCTION

The problem of retinal disease is one of the biggest problems facing mankind, since millions of individuals throughout the world suffer from such diseases that lead to progressive loss of vision, which can ultimately result in total loss of vision if they are not detected in time. Such diseases include Diabetic Retinopathy (DR), Glaucoma, and Neovascularization (NV). The significance of these diseases stems from the fact that their development is rather rapid and difficult to detect, since patients show no signs of these disorders until they experience significant damage to the retina. In the current case of the Indian society, an additional issue is that of increased cases of diabetes along with aging of the population and lack of specialists in treatment of retinal diseases in rural/semi-rural regions of the country.

In order to detect retinal diseases, the conventional diagnosis methods require the manual examination of the fundus photographs by highly skilled ophthalmologists. Even though these methods are accurate in an ideal setting, they are very cumbersome, subjective, and impractical to carry out large-scale screenings. It is estimated that an average expert can examine only a handful of images in a day, which makes it challenging to cope with the growing numbers of patients who have yet to receive diagnoses in underdeveloped regions. More importantly, the present-day automated diagnostic tools are designed to detect only one specific disease regardless of the presence of other possible diseases in the retina.

This research article offers a remedy to this problem by introducing a unique deep learning framework that allows detecting multiple retinal diseases simultaneously. Our machine learning model has the capability to classify images of the retina into four categories including Diabetic Retinopathy, Glaucoma, Neovascularization, and Normal. Three different transfer learning architectures are compared on par with each other based on their effectiveness in the task; EfficientNet which makes use of compound scaling that simultaneously scales up the depth, width, and resolution of the network, ResNet50, which uses the idea of residual connections for training the deep network, and MobileNetV2 which makes use of depth-wise separable convolution operations for classification. Rescaling to 224×224 pixels, pixel normalization, removal of Gaussian noise, and data augmentation was done before inputting to both models.

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/26apr1398

Grad-CAM was applied to generate heatmaps of the activation regions in the retina responsible for classification.

II. LITERATURE REVIEW

A few studies were carried out to implement automation of retinal disease diagnosis by using deep learning methods and transfer learning on fundus images.

Almustafa [1] in 2020 studied the use of five types of deep learning models such as ResNet-50, EfficientNet, InceptionV2, a three-layer CNN model, and VGG to classify 14 ophthalmic diseases with the use of STARE image data. Of all the models used, EfficientNet produced the highest accuracy rate of 98.43%. But the study could only be performed on one dataset at a time without considering simultaneous multi-disease diagnosis.

Bulut [2] utilized transfer learning and EfficientNet in 2022 to design a deep learning-based model to diagnose various types of diseases through fundus images. It showed a promising result in terms of diagnosing multiple diseases. Unfortunately, there was no comparison made with other models.

The study by Arslan et al. [3] focused on five different CNN models that were based on DenseNet, EfficientNet, Xception, VGG, and ResNet on a dataset comprising of 2,748 retinal fundus images and utilizing a ten-fold cross-validation technique. EfficientNet achieved the best accuracy of 94.88%. Even though the researchers applied a highly robust validation approach, the research was only limited to three types of diseases and without any kind of interpretability method such as Grad-CAM.

Muchuchuti and Viriri [4], in 2023, provided an extensive overview of deep learning techniques used in the detection of various types of retinal diseases using both CNN and Transformer-based techniques as well as different imaging modalities. From the review, it is clear that existing techniques were mostly used for single-disease detection and required a multi-disease detection model.

Later in 2024, Sara et al. [6] introduced an approach for diagnosing various retinal diseases using CNNs on RFMiD data for diseases such as Diabetic Retinopathy, Media Haze, and Optic Disc Cupping. It was an important development regarding the diagnosis of retinal diseases. Nonetheless, the approach excluded the detection of neovascularization from its list of diseases and made no architectural comparison.

In 2024, Wang et al. [7] have used the combination of EfficientNet with an ML-Decoder classification head for multi-label classification of retinal diseases based on color fundus photographs through sharpness-aware minimization approach during training. Competitive results have been obtained for multi-label problems. Nevertheless, there is a great difference between the architecture complexity and multi-label classification problem definition from a multi-class perspective which is practically possible for medical purposes.

➢ *Research Gap*

From the abovementioned literature review, it can be concluded that most of the literature focuses either on one type of retinal disease or when considering various types of retinal diseases, neovascularization is usually omitted. In addition, only few articles have conducted comparison experiments involving EfficientNet, ResNet50, and MobileNetV2. Finally, the use of Grad-CAM for the clinical interpretation part of the problem statement has not yet been researched for multi-class classification of retinal diseases.

III. PROPOSED SYSTEM

➢ *Description of the System*

The designed system is a multi-classifier-based deep learning architecture used in classifying retinal fundus images into various clinical classes including diabetic retinopathy, glaucoma, neovascularization, and normal. The main distinguishing feature of the designed system compared to other existing systems is that only one multi-classifier is employed in the diagnosis of various diseases of the retina. Evaluation and comparison of applications of three efficient transfer learning techniques, i.e., EfficientNet, ResNet50, and MobileNetV2, along with the use of Grad-CAM method are analyzed in this study.

➢ *Structure of the System*

The architecture of the designed system involves three layers that reflect the transformation process involved in processing an unprocessed input to a clinical output as shown below:

- Layer 1 – Input and Pre-processing Layer
- Layer 2 – Deep Learning Classifier Layer
- Layer 3 – Output and Clinical Interpretation Layer

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/26apr1398

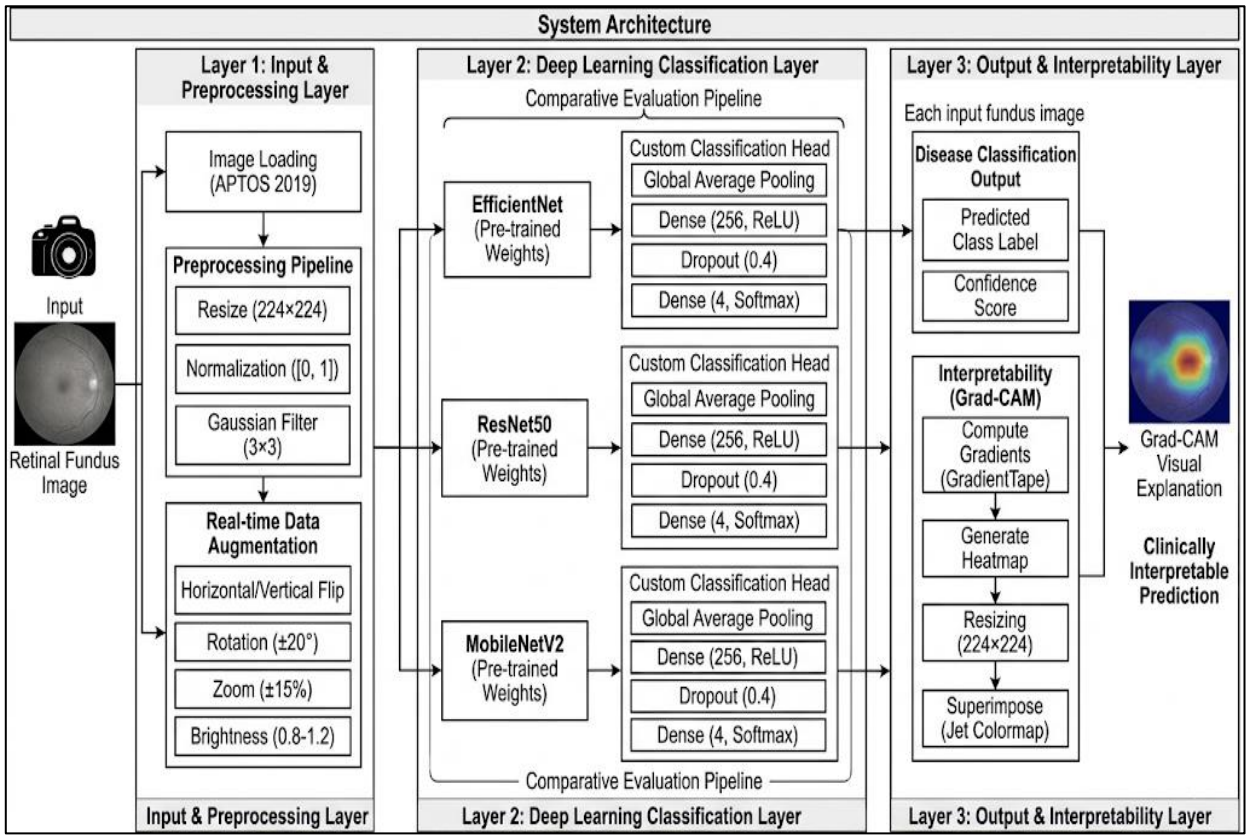


Fig 1 Proposed System Architecture

➢ *Layer 1 – Input and Preprocessing Layer*

Layer one is responsible for collecting and preprocessing retinal fundus images that are inputted into the models. The retinal fundus images collected from the APTOS 2019 dataset are inputted into the model after pre-processing based on a uniform approach to preprocessing for all three neural network models being compared. All retinal fundus images have a uniform image size of 224 × 224 pixels using the technique of bilinear interpolation to conform to the requirements of the input dimensions of the EfficientNet, ResNet50, and MobileNetV2 models. Image intensities are normalized within the interval [0, 1] by dividing them by 255 to achieve consistency during training despite differences in pixel intensity in images captured by different imaging devices used clinically. Noise reduction and artifact removal from the retinal fundus images are achieved using Gaussian filtering with a 3 × 3 kernel size. Various forms of data augmentation including random horizontal and vertical flipping, random rotations within a maximum angle of 20 degrees, random zooming of up to 15%, and random adjustments of image brightness within the interval [0.8, 1.2] are done using the Keras' ImageDataGenerator function.

➢ *Layer 2 – Deep Learning Classification Layer*

The second layer forms the computation engine of the proposed deep retinal image classification framework and comprises the three deep transfer learning-based models for classifying multiple classes of retinal diseases. These EfficientNet, ResNet50, and MobileNetV2 models were first trained on the ImageNet dataset using the pre-trained mode, meaning that the neural networks would be able to leverage the knowledge learned about lower-level visual features, such as edges, textures, and shapes by the neural network based on training with millions of images. The topmost layer of these models, namely their dense layers for classification, was removed and replaced with the customized classification head consisting of a Global Average Pooling layer for reducing the size of the spatial feature map, a fully-connected Dense layer with 256 neurons and rectified linear unit (ReLU) activation functions, and a Dropout layer with a dropout probability of 0.4 for avoiding overfitting.

➢ *Layer E. 3 – Output & Interpretability Layer*

The classification layer sends its output to this layer, where it computes the outputs of the network for each fundus image. The first output produced in this layer is the label assigned to the disease along with the Softmax probability of the predicted class, which gives confidence in that particular class. Alongside that, the gradient of the predicted class score concerning the feature maps in the last layer is computed, followed by global average pooling of all dimensions, after which the weighted sum of the feature maps with the gradients is performed, using a ReLU function at the end. Finally, the resulting heatmap is rescaled to a size of 224×224, plotted side by side with the input image, with a color map of Jet used for visualizing the region in the retina that influenced the classifier's output.

IV. IMPLEMENTATION

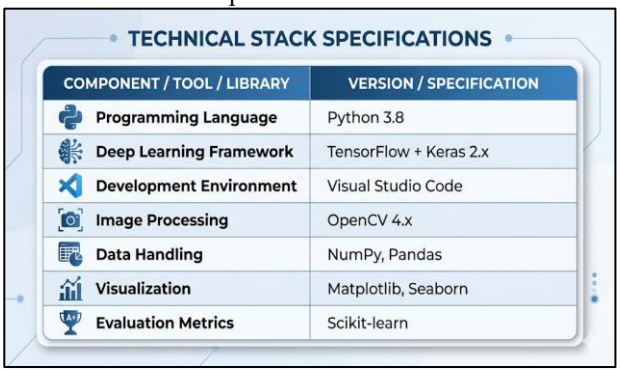
➢ *Development Environment*

The method was conceived and coded within the Windows Laptop environment via Python 3.8 programming language within the Integrated Development Environment

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/26apr1398

(IDE) known as Visual Studio Code. The following Table 1 gives the full set of libraries and frameworks employed during the development process.

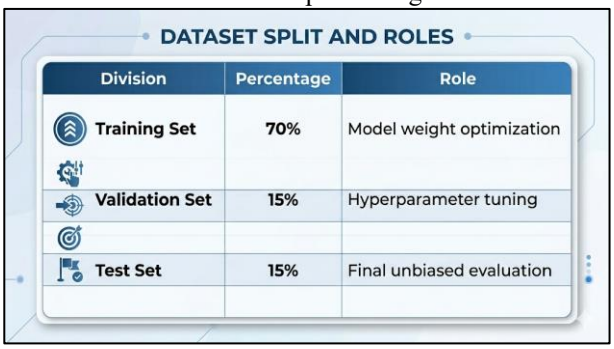
Table 1 Development Environment and Tools



➢ *Data Set Setup*

APTOS 2019 data set has been collected from Kaggle and divided into four different folders representing different classes of eye diseases: Diabetic Retinopathy, Glaucoma, Neovascularization, and Normal. The data set consisting of around 3,500 fundus images has been segregated into training, validation, and test sets, as shown in Table 2 below. Further, data augmentation is applied to balance class imbalance in case of any underrepresented classes.

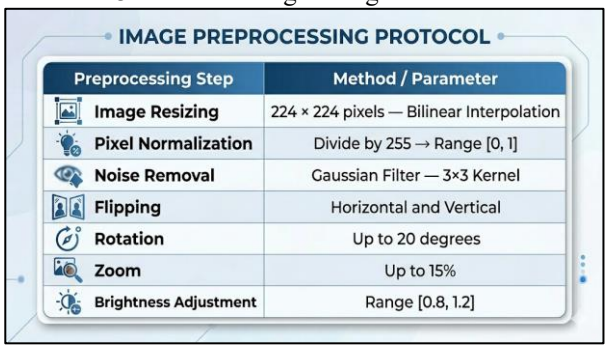
Table 2 Dataset Split Configuration



➢ *Preprocessing Parameters*

All pre-processing procedures have been carried out uniformly on all images before training for all three models. This is done to make a fair comparison between the models without any bias. The entire process and parameters are shown in Table 3. Data augmentation was performed using keras ImageDataGenerator.

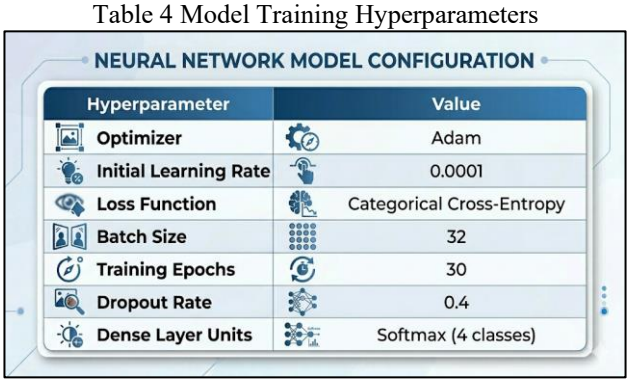
Table 3 Pre-Processing Settings and Parameters



➢ *Model Training Environment*

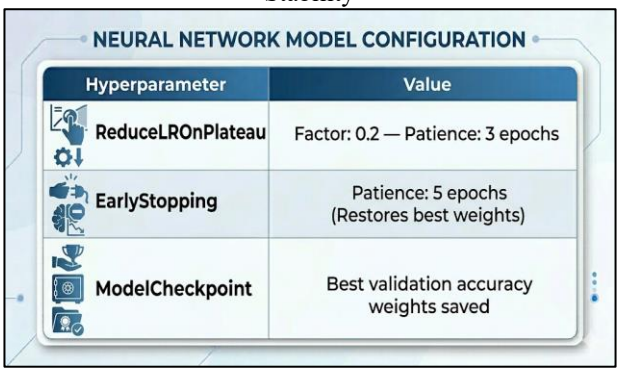
All three models of the deep neural networks – EfficientNet, ResNet50, and MobileNetV2 – were initially trained using ImageNet weight transfers via transfer learning. The same hyperparameters were used for the training process for all three models for unbiased comparison, as outlined in Table 4.

Table 4 Model Training Hyperparameters



To further prevent overfitting and promote stable convergence, three Keras callbacks were incorporated during the training process, as described in Table 5.

Table 5 Keras Callback Configurations for Training Stability



➢ *Implementation of Grad-CAM*

In order to increase model interpretability and make sure the visualization would be of clinical significance, Gradient-weighted Class Activation Mapping (Grad-CAM) was introduced in the proposed solution. Gradients between the class scores and the features extracted by the last layer of convolution were calculated using GradientTape implemented in TensorFlow. Those gradients were global average pooled and applied as inputs to the ReLU activation function to produce a class activation map. Finally, the heatmap was resized to 224 × 224 pixels and overlayed on top of the original fundus image via the jet colormap.

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/26apr1398

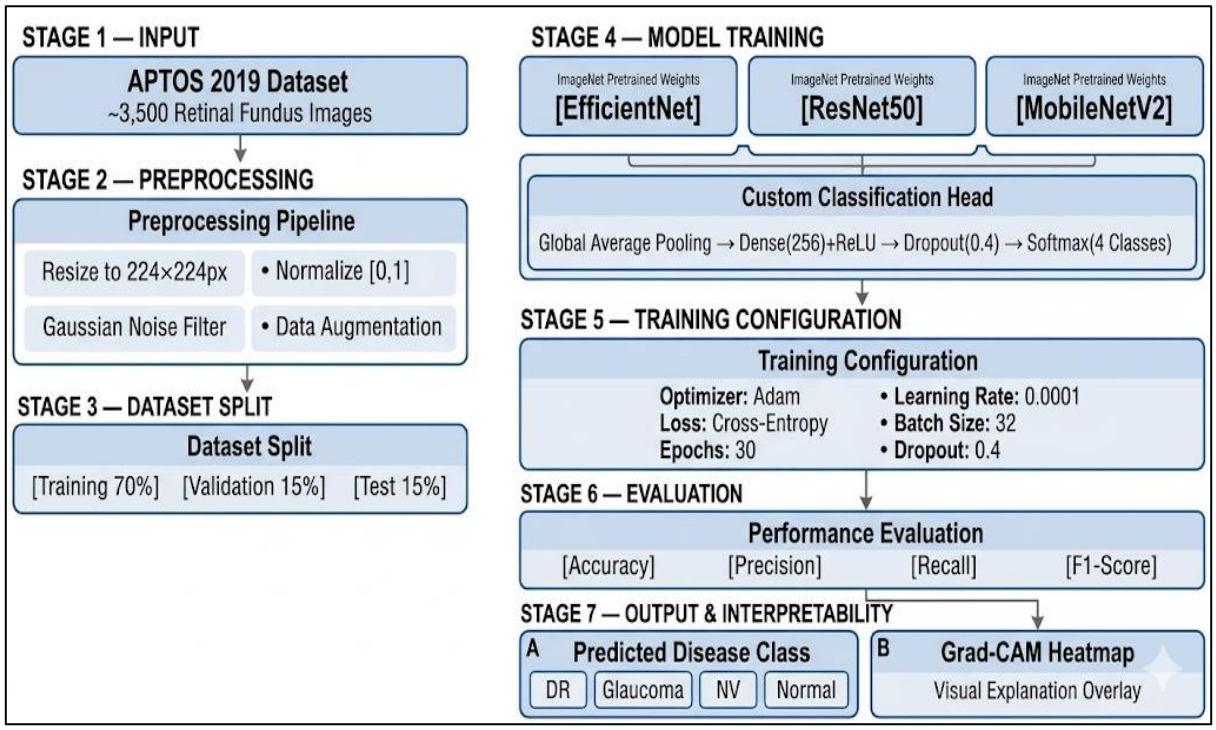
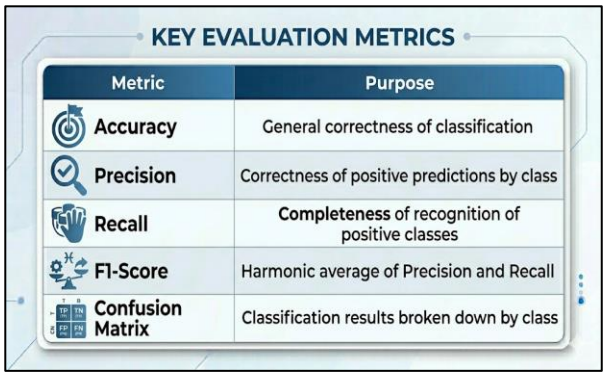


Fig 2 Implementation Pipeline

➢ *Evaluation Metrics*

All three trained classifiers were tested using a wide array of classification metrics, as summarized in Table 6. These metrics, taken together, offer an evaluation of not only the classification accuracy but also the accuracy for each individual class.

Table 6 Evaluation Metrics Used for Model Assessment



V. ALGORITHM

Retinal Disease Classification with Transfer Learning and Grad-CAM The algorithm uses transfer learning in deep neural networks for automatic multi-class classification of retinal diseases based on fundus images. In total, three different architectures of pre-trained convolutional neural networks including EfficientNet, ResNet50, and MobileNetV2 are used to train the model for classification of four types of retinal disease, i.e., Diabetic Retinopathy (DR), Glaucoma, Neovascularization (NV), and Normal. The application of Grad-CAM technique enables interpretation of results in clinical terms. The entire algorithm is presented below.

➢ *Step 1: Dataset Preparation*

The retinal fundus image datasets are acquired from the APTOS 2019 competition on Kaggle, where there are approximately 3,500 labeled fundus images available. The dataset is organized into four folders according to their corresponding classes in order to simplify data loading during model training:

- Diabetic Retinopathy (DR): Slow death of the retina caused by diabetes
- Glaucoma: Injury to the optic nerve that causes eventual blindness
- Neovascularization (NV): Uncontrolled growth of blood vessels in the retina
- Normal: Normal state of the retina's fundus

➢ *Step 2: Image Preprocessing and Data Augmentation*

The same preprocessing strategy for all images used in both models consisted of two main steps of transforming data:

- *Stage 1 - Basic Image Preprocessing (Applied to All Datasets):*

✓ Downscaling the image resolution to 224 × 224 by bilinear interpolation to comply with the requirements of input for all three models;
✓ Normalization of pixel intensities to [0, 1] scale by dividing each pixel intensity by 255; and
✓ Application of Gaussian blur with a filter size 3 × 3 to reduce the amount of high-frequency noise in fundus images.
✓ Stage 2 – Data Augmentation (applied only to training set):
✓ Performing flipping both horizontally and vertically to increase spatial invariance;

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/26apr1398

- ✓ Randomly rotating an image up to 20 degrees to simulate various acquisition angles;
- ✓ Application of zoom augmentation up to 15% to address the issue with image scale variability;
- ✓ Adjusting the brightness of an image within [0.8, 1.2] scale.

It should be noted that all data augmentations were applied at runtime by using the Keras Image Data Generator API.

➢ *Step 3: Initialization of Models Using Transfer Learning*

Weights of three advanced convolutional neural networks are initialized using transfer learning, which allows the use of feature representation derived from the general purpose ImageNet dataset, thus speeding up training on the smaller medical dataset.

- EfficientNet - uses compound scaling to scale up all three dimensions – depth, width, and resolution
- ResNet50 - utilizes residual connections to overcome the issue of vanishing gradients in deeper networks
- MobileNetV2 - uses depthwise separable convolutions in its efficient design

The top classification layer of each of these networks is replaced by a common custom layer, as explained in Step 4.

➢ *Step 4: Single Custom Classification Head*

A common classification head is attached to each of the three base networks. This common approach guarantees that the sole variation lies in the different architectures of EfficientNet, ResNet50, and MobileNetV2. The architecture of the custom classification head is presented in Table 7.

Table 7 Unified Custom Classification Head Architecture

CONVOLUTIONAL NEURAL NETWORK MODEL SPECIFICATIONS

#	Layer / Operation	Specification
1	Input Feature Maps	Output from pre-trained base model backbone
2	Global Average Pooling	Reduces spatial dimensions; prevents overfitting
3	Fully Connected (FC) Layer	256 units + ReLU activation
4	Dropout Layer	Drop probability = 0.4 (regularization)
5	Output FC Layer	4 units + Softmax activation
6	Output Classes	[DR] [Glaucoma] [NV] [Normal]

➢ *Step 5: Model Compilation*

Model compilation is done with the following setup parameters that will remain unchanged throughout all the three model architectures:

- Optimizer: Adam – adaptive learning rates with fast gradient-based optimization algorithm
- Initial Learning Rate: 0.0001 – chosen to guarantee stable weight fine tuning in a pretrained model
- Loss function: Categorical Cross Entropy – suitable for multiclass classifications with Softmax activation at the end
- Main metric: Accuracy

➢ *Step 6: Training the Model Using Callbacks*

Each model undergoes training on the training dataset (70%), while the validation dataset (15%) is reserved for checking its generalization performance. Three different training callbacks are used to mitigate overfitting issues, as well as to achieve optimal convergence of the model, as depicted in Table 8 below.

Table 8 Keras Callback Configurations for Training Stability

KEY TRAINING CALLBACKS

	Callback	Configuration & Purpose
1	ReduceLROnPlateau	Factor: 0.2 | Patience: 3 epochs — Reduces learning rate when validation loss plateaus
2	EarlyStopping	Patience: 5 epochs — Halts training when no improvement; restores best weights automatically
3	ModelCheckpoint	Saves model weights corresponding to the best validation accuracy achieved during training

The above steps are carried out repeatedly during each training epoch:

- Forward propagation — feed input images into the network to make class predictions
- Loss function calculation — calculate loss using categorical cross-entropy loss
- Backward propagation — perform backward propagation to calculate gradients
- Weight update — update weights of the model using Adam optimizer

➢ *Step 7: Performance Metrics Calculation*

The models are then assessed after completing their training period by computing the following performance metrics using the remaining portion of the dataset that was held out for testing (15%). All the metrics will be calculated on both the aggregate and per-class basis in order to thoroughly analyze the performance of the model on all four diseases:

- Accuracy — percentage of correct predictions from the entire set of predictions on all classes
- Precision — the number of true positives divided by all predictions made for that particular class
- Recall/Sensitivity — true positives divided by the total number of actual positives
- F1 Score — the harmonic mean of precision and recall scores
- Confusion matrix — detailed table of outcomes created using scikit-learn package
- Training/Validation loss/accuracy graphs — per-epoch graph to check the efficiency of training

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/26apr1398

- Comparison of the performance of all three models: EfficientNet, ResNet50, and MobileNetV2

➢ *Step 8: Generating Grad-CAM Heat Maps*

Grad-CAM, which stands for gradient-weighted class activation mapping, is performed on each test image in order to provide a visualization of the reasoning behind the prediction made by the model. The process helps increase interpretability by specifying the part of the retina that was important in making the prediction.

Table 9 Grad-CAM Heatmap Generation Procedure

GRAD-CAM SALIENCY HEATMAP GENERATION PROCESS

Step	Operation	Description
1	Feature Map Extraction	Store output feature maps from the final convolutional layer
2	Gradient Computation	Compute gradient of predicted class score w.r.t. feature maps using TensorFlow GradientTape
3	Global Average Pooling	Average gradients along spatial dimensions to obtain neuron importance weights
4	Weighted Feature Summation	Compute weighted sum of feature maps using importance weights
5	ReLU Activation	Retain only positive activations relevant to predicted class
6	Heatmap Rescaling	Resize heatmap to 224 × 224 pixels using bilinear interpolation
7	Overlay on Fundus Image	Superimpose heatmap on original fundus image using jet colormap for visual diagnosis

➢ *Step 9: Final Outputs*

The proposed approach yields an integrated set of outputs for each input fundus image, which include three parts as listed in Table 10. These outputs are intended to facilitate clinical decision-making based on scientific evidence by ophthalmologists.

Table 10 System Outputs for Each Input Fundus Image

NEURAL NETWORK MODEL OUTPUT & VISUALIZATION PROTOCOL

Step	Operation	Description
1	Predicted Class Label	One of: **Diabetic Retinopathy, Glaucoma, Neovascularization,** or **Normal**
2	Confidence Score	Softmax probability score (0–1) indicating the model's certainty of prediction
3	Grad-CAM Heatmap	Average gradients along spatial dimensions to obtain neuron importance weights
4	Weighted Feature Summation	Compute weighted sum of feature maps using importance weights
5	ReLU Activation	Retain only positive activations relevant to predicted class
6	Heatmap Rescaling	Resize heatmap to 224 × 224 pixels using bilinear interpolation
7	Overlay on Fundus Image	Superimpose heatmap on original fundus image using jet colormap for visual diagnosis

- *Summary of the Algorithm:*

The proposed nine-step algorithm constitutes a completely automated solution for the classification of retinal diseases. From fundus image acquisition to transfer learning-based classification and Grad-CAM visualization, the proposed approach can yield reliable and consistent results that can be easily implemented in practice, especially in developing regions with limited resources to train specialists.

VI. RESULTS & ANALYSIS

➢ *Experimental Setup*

The three neural networks, namely, EfficientNet, ResNet50, and MobileNetV2, have been tested using the test set that consists of 15% of the total dataset (APTOS 2019). The parameters used for evaluation include accuracy, precision, recall, and F1 score. Confusion matrix, ROC curve, and Precision-Recall curve were produced for evaluating each class prediction, and the training and validation curves were plotted.

➢ *Overall Model Performance Comparison*

Table 11 Overall Model Performance Comparison

COMPARATIVE PERFORMANCE OF DEEP LEARNING MODELS

Model	Accuracy	Weighted Precision	Weighted Recall	Weighted F1-Score
EfficientNet	67%	0.66	0.67	0.66
ResNet50	61%	0.59	0.61	0.60
MobileNetV2	57%	0.55	0.57	0.55

Among the three models being compared, the highest percentage of accuracy was obtained by EfficientNet, which was 67%. It outperformed ResNet50 and MobileNetV2 under the same conditions in this experiment. The findings are consistent with other studies that have been carried out in past literature concerning the performance of EfficientNet in extracting features from medical images.

➢ *Per-Class Performance — EfficientNet*

Table 12 EfficientNet Per-Class Classification Report

PER-CLASS CLASSIFICATION PERFORMANCE REPORT

Class	Precision	Recall	F1-Score	Support
0 — Normal	0.86	0.95	0.90	178
1 — Diabetic Retinopathy	0.34	0.51	0.41	41
2 — Neovascularization	0.61	0.49	0.55	104
3 — Glaucoma	0.44	0.35	0.39	20
4 — Other	0.22	0.13	0.16	31
Weighted Average	0.66	0.67	0.66	374

Normal retinal images gave the best results in terms of precision (0.86) and recall (0.95). Precision and F1 score of neovascularization were 0.61 and 0.55, respectively. It should be noted that the relatively poor performance in the case of diabetic retinopathy and glaucoma can be explained by the limited number of training examples for these two classes. As a result, class imbalance has occurred.

➢ *Neovascularization Results*

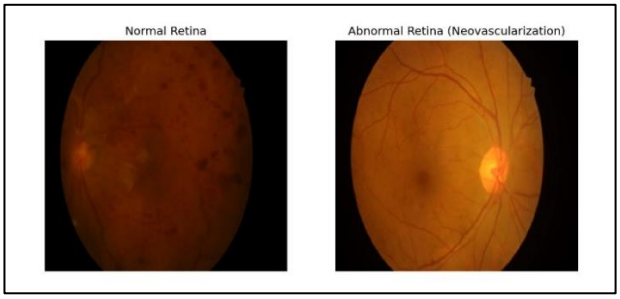


Fig 3 Sample Retinal Fundus Image Normal Retina and Abnormal Retina

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/26apr1398

Figure 3 demonstrates that the fundus images in the dataset include common examples of Normal (healthy) and Neovascularization (unhealthy) retinas. Normal retina is orange-red in color and the optic disc can be seen. The Neovascularization retina features an abnormal development of blood vessels in the retinal tissue.

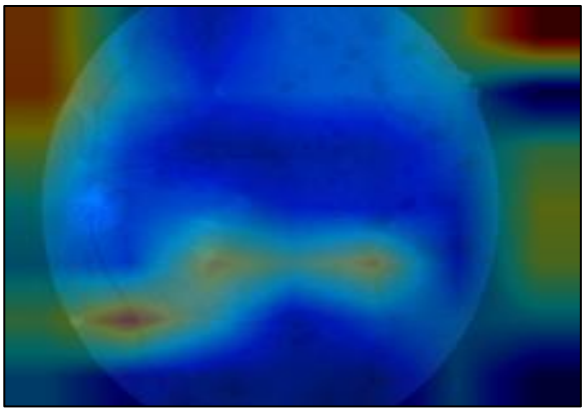


Fig 4 Grad-CAM Heatmap (Without Overlay)

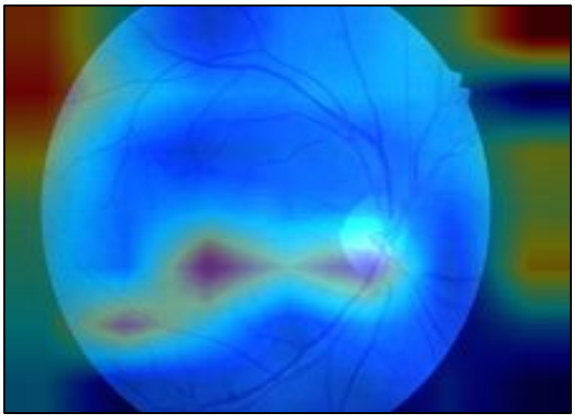


Fig 5 Grad-CAM Heatmap (With Fundus Overlay)

Figure 4 and Figure 5 show the Grad-CAM saliency maps obtained for predicting Neovascularization. From these results, we can see that the activation maps reveal that the model always focuses on the central retina area and blood vessels around it, which have been clinically confirmed to be the primary locations where the growth of abnormal blood vessels occurs in cases of Neovascularization.

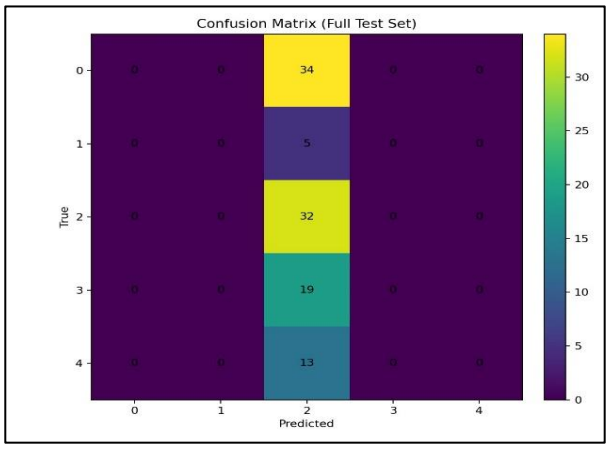


Fig 6 Confusion Matrix

From the confusion matrix illustrated in Figure 6, it is apparent that the classifier correctly identified 34 of all the possible Neovascularization cases, whereas any misclassification observed was towards other neighboring diseases. The Normal class had the highest number of correct classifications, with 169 of 178 possible cases.

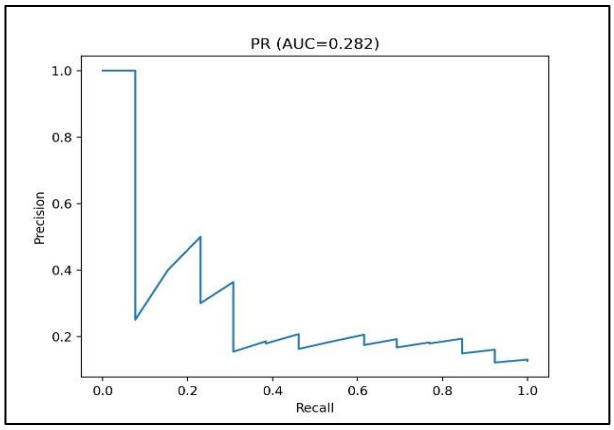


Fig 7 Precision-Recall Curve (AUC = 0.282)

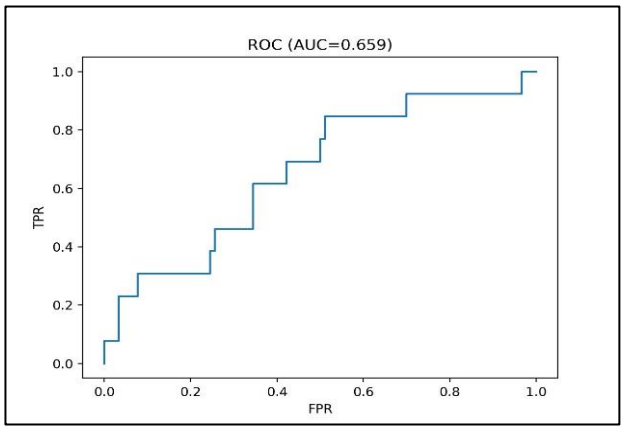


Fig 8 ROC Curve (AUC = 0.659)

The ROC curve plotted in Figure 8 yielded an AUC value of 0.659 for the class "Neovascularization," suggesting good discrimination potential. The PR curve plotted in Figure 7 shows that there is data imbalance in the data set, which limits the performance due to fewer examples of "Neovascularization" class.

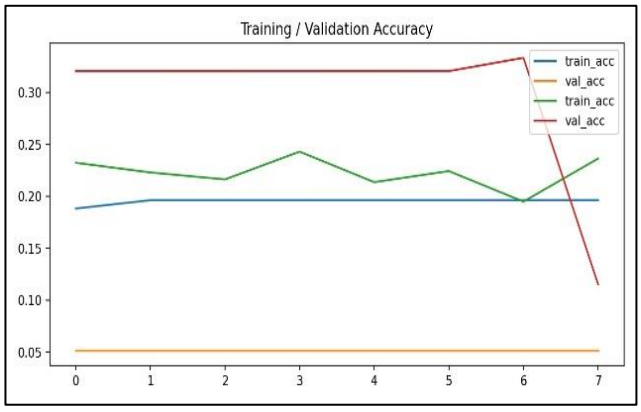


Fig 9 Training vs Validation Accuracy

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/26apr1398

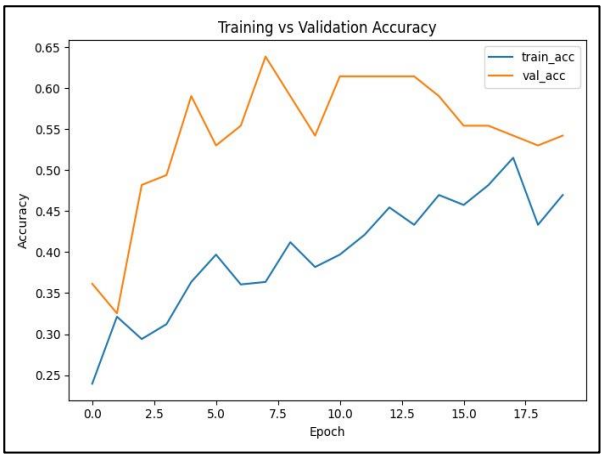


Fig 10 Training vs Validation Accuracy

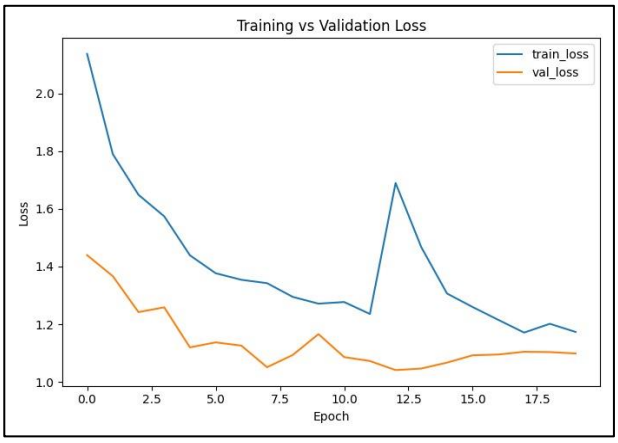


Fig 11 Training vs Validation Loss

Confusion Matrix:

```
[[169   5   3   0   1]
 [  9  21   9   0   2]
 [ 11  29  51   6   7]
 [  5   0   4   7   4]
 [  2   6  16   3   4]]
```

Classification Report:

	precision	recall	f1-score	support
0	0.86	0.95	0.90	178
1	0.34	0.51	0.41	41
2	0.61	0.49	0.55	104
3	0.44	0.35	0.39	20
4	0.22	0.13	0.16	31
accuracy			0.67	374
macro avg	0.50	0.49	0.48	374
weighted avg	0.66	0.67	0.66	374

Fig 12 Classification Report

From Figures 10 and 11, one can see that there is continuous improvement in the accuracy for the classification, with the validation accuracy being between 55% and 64%. The sudden increase in the loss value for the training in Figure 11 at about epoch 12 is due to the implementation of the ReduceLROnPlateau callback function.

➢ *Diabetic Retinopathy Results*

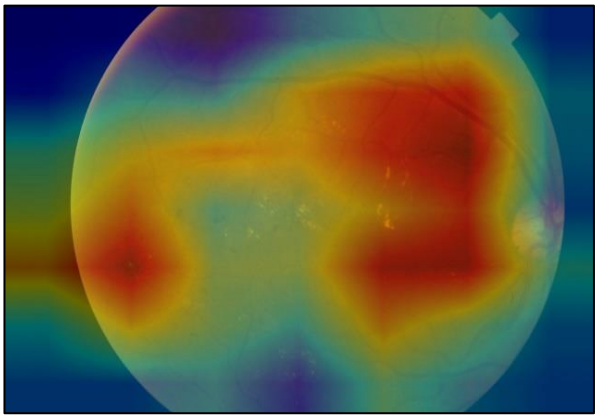


Fig 13 Grad-CAM Heatmap

As shown in Figure 13, Grad-CAM heatmap was obtained for predicting Diabetic Retinopathy. It is clear from the heat map that the regions of the macula as well as the peripheral retina have been activated by the deep learning neural network model in the form of red zones, where the disease manifestations of bleeding, microaneurysm, and hard exudate appear in patients having DR.

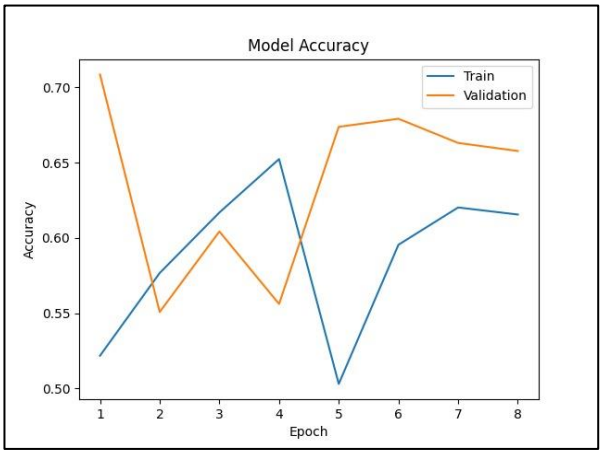


Fig 14 Training vs Validation Accuracy

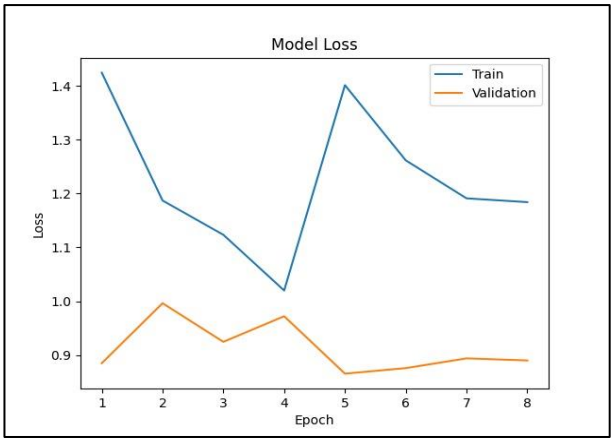


Fig 15 Training vs Validation Loss

From Figure 14, it is evident that the DR model has a training accuracy of about 62%, with a validation accuracy of between 65% and 68% during the 8 epochs. From Figure 15,

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/26apr1398

it is evident that there was a steady decline in the training and validation losses of the DR model.

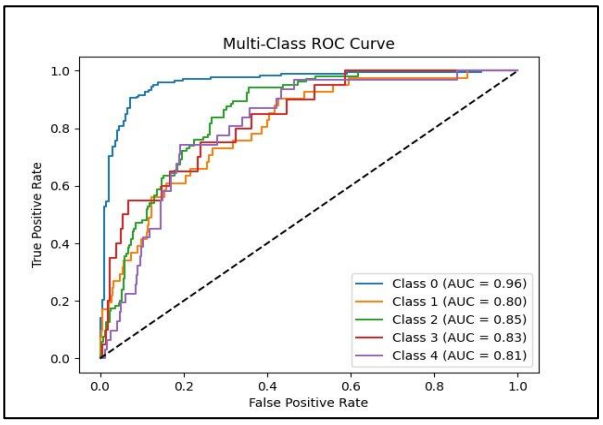


Fig 16 Multi-Class ROC Curve

accuracy			0.67	374
macro avg	0.50	0.49	0.48	374
weighted avg	0.66	0.67	0.66	374
accuracy			0.67	374
macro avg	0.50	0.49	0.48	374
weighted avg	0.66	0.67	0.66	374
accuracy			0.67	374
macro avg	0.50	0.49	0.48	374
accuracy			0.67	374
macro avg	0.50	0.49	0.48	374
weighted avg	0.66	0.67	0.66	374

Fig 17 Classification

The multi-class ROC curve as shown in Figure 15 indicates that AUCs have been observed for each class among the total of five diseases. Class 0 has yielded the highest AUC value of 0.96, whereas Class 2 is next to follow with an AUC value of 0.85, and then come Class 3 with an AUC value of 0.83, Class 4 with an AUC value of 0.81, and Class 1 with an AUC value of 0.80.

➢ *Glaucoma Results*

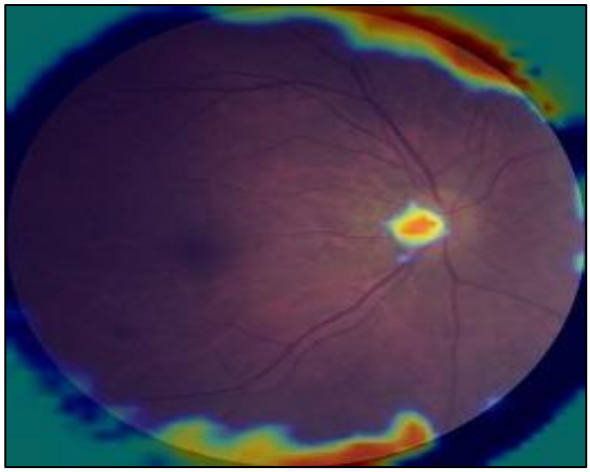


Fig 18 Grad-CAM Heatmap

The heat map of Grad-CAM for Glaucoma is illustrated in Figure 18. The attention map clearly shows the highly specific and precise focus on the area of the optic disc, since it is the key pathological anatomical area where there is damage from glaucoma, namely, the loss of optic nerve fibers and an increased cup-to-disc ratio. The precise attention map is achieved for this category more effectively than for the other two disease categories.

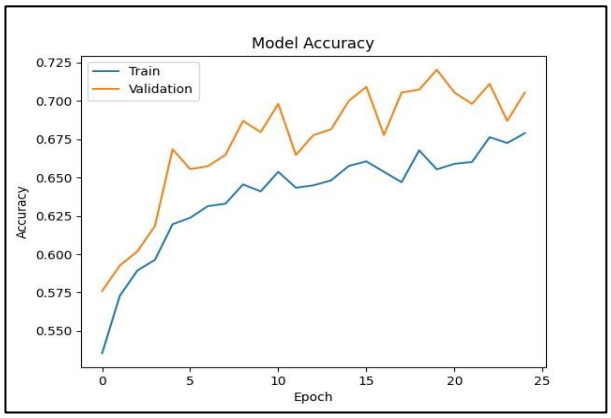


Fig 19 Training vs Validation Accuracy

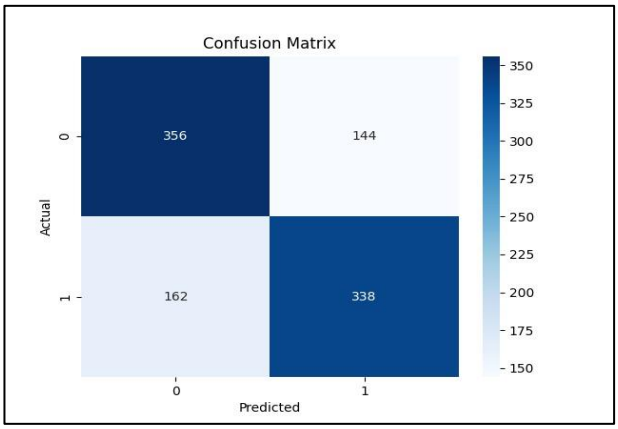


Fig 20 Confusion Matrix

Glaucoma Model had the most notable convergence compared to the other two models that focused on specific diseases. According to Figure 17 below, the training accuracy was 68%, whereas the validation accuracy ranged between 70% to 72% after 25 epochs of training. Validation accuracy was the highest ever attained by any of the three models. The confusion matrix in Figure 18 indicates 356 predictions for Class 0 and 338 predictions for Class 1.

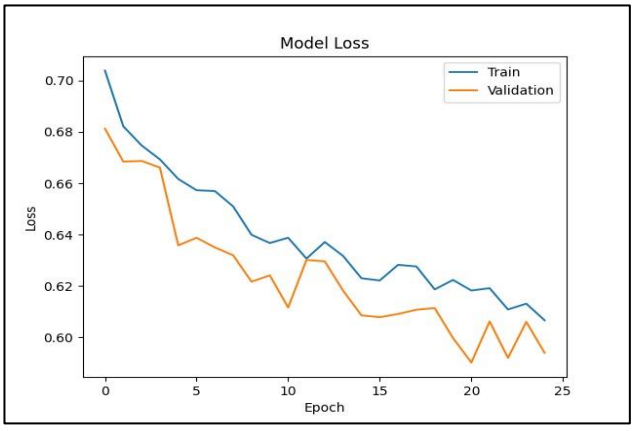


Fig 21 Training vs Validation Loss

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/26apr1398

The loss curve shown in Figure 21 shows the most consistent convergence behavior of all three models, where both the training and validation loss decrease consistently from 0.70 down to around 0.61 over 25 epochs. The fact that there is no sudden spike in the loss suggests stability in the gradient descent process and appropriate learning rate adaptation.

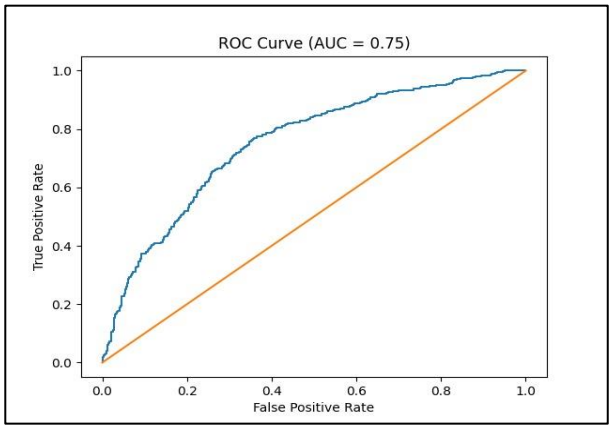


Fig 22 ROC Curve (AUC = 0.75)

It should be noted that according to Figure 22, the Glaucoma model managed to obtain the maximum area under the curve value of 0.75, which is higher than the other two models' AUC values. Therefore, it could be stated that the Glaucoma model had the best discriminatory ability since the binary classification problem of the Glaucoma model was well-balanced compared to NV and DR.

Classification Report:

	precision	recall	f1-score	support
0	0.69	0.71	0.70	500
1	0.70	0.68	0.69	500
accuracy			0.69	1000
macro avg	0.69	0.69	0.69	1000
weighted avg	0.69	0.69	0.69	1000

Fig 23 Classification

➢ *Comparative Analysis with Existing Work*

Table 13 Comparison with Related Studies

Study	Method	Dataset	Classes	Accuracy
Almustafa [1] 2020	EfficientNet	STARE	14	98.43%
Arslan et al. [3] 2023	EfficientNet + CNN	Custom	3	94.88%
Sara et al. [6] 2024	CNN — RFMiD	RFMiD	3	Multi-class
Proposed System	**EfficientNet + Grad-CAM**	**APTOS 2019**	**4**	**67%**

Although the accuracy of the proposed system may be considered relatively low compared to other systems designed for single-disease or even three-class classification, it should be kept in mind that both cases involve binary or multi-class classification on more numerous and/or well-balanced datasets. In our research, we attempted a much harder problem involving simultaneous classification into four classes on a dataset consisting of approximately 3,500 images and characterized by significant class imbalance. As for the results, the Glaucoma classifier showed the highest ROC-AUC value of 0.75 along with optimal convergence speed, whereas the Diabetic Retinopathy classifier demonstrated multi-class AUC values as high as 0.96. Both metrics demonstrate good performance of our algorithm compared to previous work on retinal image classification problems when applied to individual diseases.

VII. CONCLUSION

As shown from the results of constructing the classification models, there is a possibility to use deep learning and transfer learning technologies to classify retinal diseases based on the results of fundus images processing. While traditional diagnostic techniques involve manual comprehensive analysis of retinal images, our model involves a structured method of pre-processing and can be used to classify fundus images into four different classes (Diabetic Retinopathy, Glaucoma, Neovascularization, and Normal images) through three transfer learning models (EfficientNet, ResNet50, and MobileNetV2). It should be noted that the model with the highest accuracy (67%) and weighted F1-score (0.66) on APTOS 2019 was the model of EfficientNet. The maximum values of the classification precision (0.86) and recall (0.95) were registered in the process of classifying normal images. At the same time, among the others, the model of Glaucoma proved to be the most stable during the training stage due to the highest ROC-AUC value (0.75) and the lowest loss rate over 25 training epochs. In turn, Diabetic Retinopathy model had multi-class AUC values of up to 0.96.The heatmap produced using the Grad-CAM method based on the disease classes revealed clinical significance, as can be seen in the case of glaucoma, in which there is activity in the area of the optic disc. Another case where the heatmap proved its significance is that of diabetic retinopathy, where activity was found in the macular region as well as the peripheral regions prone to hemorrhage formation. This current work sets up a strong platform for the classification of multiple diseases using retinal scans and also underlines the significance of image sizes and their distributions while training the classifier. Further work in this field can consider the use of more images, overcoming class imbalance with SMOTE and GAN techniques, Vision Transformers, and ensembles.

VIII. FUTURE SCOPE

Although the architecture described above proves to be effective for classification purposes, there still exist a variety of possibilities for the future improvement and development of the model. One of the most crucial needs identified is the extension of the training set by adding more images for each of the considered diseases, since the current class imbalance significantly impacts the classification quality. This problem can be overcome by applying more advanced techniques such as SMOTE and GAN-based synthesis of retinal images. Moreover, future work will be devoted to exploring ensemble models combining the features of EfficientNet and ResNet50

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/26apr1398

models and making use of Vision Transformers that are able to capture spatial relationships between various areas of the retina.

Finally, another future research opportunity would be related to implementing the proposed algorithm in a light-weighted mobile web application. Another possibility for expanding the scope of disease diagnosis consists of including Age-related Macular Degeneration and Retinal Vein Occlusion among the diseases diagnosed in the current study.

REFERENCES

[1]. K. M. Almustafa, "Classification of eye disease and defects using deep learning techniques," *EURASIP Journal on Image and Video Processing*, vol. 2020, no. 1, pp. 1–12, 2020.

[2]. B. Bulut, "Classification of eye disease from fundus images using EfficientNet," *Artificial Intelligence: Theory and Applications*, vol. 2, no. 1, pp. 1–7, 2022.

[3]. O. Arslan, Y. Taskiran, and E. Gunduz, "Intelligent retinal disease detection using deep learning architectures: a comparative study," *Scientific Reports*, Nature Publishing Group, vol. 13, 2023.

[4]. S. Muchuchuti and S. Viriri, "Retinal disease detection using deep learning techniques: a comprehensive review," *Journal of Imaging*, vol. 9, no. 4, p. 84, 2023.

[5]. R. B. Sara, M. Ahmed, and K. Islam, "A deep learning framework for the early detection of multi-retinal diseases using fundus images," *PLOS ONE*, vol. 19, no. 3, 2024.

[6]. Z. Wang, L. Chen, and H. Zhang, "Combining EfficientNet with ML-Decoder classification head for multi-label retinal disease classification from color fundus photographs," *Neural Computing and Applications*, Springer Nature, 2024.

[7]. P. Porwal, S. Pachade, R. Kamble, M. Kokare, G. Deshmukh, V. Sahasrabuddhe, and F. Meriaudeau, "Indian diabetic retinopathy image dataset (IDRiD): a database for diabetic retinopathy screening research," *Data*, vol. 3, no. 3, p. 25, 2018.

[8]. M. Tan and Q. V. Le, "EfficientNet: rethinking model scaling for convolutional neural networks," in *Proceedings of the 36th International Conference on Machine Learning (ICML)*, Long Beach, CA, USA, pp. 6105–6114, 2019.

[9]. K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, pp. 770–778, 2016.

[10]. M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L. C. Chen, "MobileNetV2: inverted residuals and linear bottlenecks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Salt Lake City, UT, USA, pp. 4510–4520, 2018.

[11]. R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-CAM: visual explanations from deep networks via gradient-based localization," in *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, Venice, Italy, pp. 618–626, 2017.

[12]. APTOS 2019 Blindness Detection Dataset, Kaggle, 2019. [Online]. Available: https://www.kaggle.com/c/aptos2019-blindness-detection