International Journal of Environmental Sciences
ISSN: 2229-7359
Vol. 11 No. 20s, 2025
https://www.theaspd.com/ijes.php

Deep Learning Methods For Classifying Multiple Retinal Diseases Using Fundus Images

C Sharmila Suttur[1*], Poornima U[2], Shreya Kulkarni[3]

[1*]Department of Electronics and Communication Engineering, Assistant Professor, M S Ramaiah Institute of Technology, Bangalore, India
[2]Department of Electronics and Communication Engineering, Post Graduate Student, M S Ramaiah Institute of Technology, Bangalore, India
[3]Department of Electronics and Communication Engineering, Post Graduate Student, M S Ramaiah Institute of Technology, Bangalore, India
*Corresponding Author: C Sharmila Suttur
*Department of Electronics and Communication Engineering, Assistant Professor Ramaiah Institute of Technology, Vidya soudha, MSR Nagar, MSRIT Post, Bangalore-560054, India Email: sharmila.c@msrit.edu

ABSTRACT

Retinal diseases such as Diabetic Retinopathy (DR), Age-Related Macular Degeneration (ARMD), Optic Disc Cupping (ODC), Glaucoma, and Myopia are leading causes of vision impairment and blindness globally. Early detection through automated screening using deep learning is crucial for effective treatment and vision preservation. This study investigates the application of deep learning methods for the multi-class classification of retinal diseases using fundus images. Seven models—VGG16, ResNet50, EfficientNetB7, DenseNet201, EfficientNetB4, ResNet152 and a Customized CNN—were evaluated on a balanced dataset of 600 images divided into six classes. Each model was assessed using performance metrics such as accuracy, sensitivity, specificity, precision, and F1-score. Among the evaluated models, the Customized CNN achieved the highest overall accuracy (91.37%) and demonstrated superior classification performance across most disease categories. The findings emphasize the potential of tailored CNN architectures in clinical screening applications and support the development of user-friendly diagnostic tools for early disease identification.

Keywords: *Retinal Diseases Classification Fundus Images Convolutional Neural Networks Deep Learning.*

1. INTRODUCTION

In recent years, the field of medical imaging has undergone significant transformation due to advances in deep learning, which enable automated, accurate, and efficient analysis of visual data. One particularly critical application is in ophthalmology, where retinal diseases such as Diabetic Retinopathy (DR), Age-Related Macular Degeneration (ARMD), Optic Disc Cupping (ODC), Glaucoma, and Myopia account for a major proportion of global vision impairment and blindness. Manual examination of fundus images by ophthalmologists, although effective, is time-consuming, subjective, and dependent on the clinician's expertise. This challenge highlights the importance of automated diagnostic tools powered by deep learning algorithms that can accurately classify multiple types of retinal pathologies from fundus images. These tools can assist clinicians by providing objective second opinions, improve diagnostic throughput, and support early detection, which is crucial to preventing irreversible damage. This study proposes and compares the effectiveness of six deep learning models—VGG16, ResNet50, EfficientNetB7, DenseNet201, EfficientNetB4, ResNet152 and a Customized CNN—for the classification of six retinal conditions. Age-Related Macular Degeneration (AMD/ARMD) – A disease affecting the central retina (macula), leading to loss of central vision [1]. It exists in two forms: dry AMD, caused by retinal thinning; and wet AMD, involving leakage from abnormal blood vessels. Diabetic Retinopathy (DR) – A complication of diabetes that damages retinal blood vessels [2]. It progresses through stages: mild non-proliferative, moderate, severe non-proliferative, and proliferative DR. Early signs include micro aneurysms and exudates, which can be identified in fundus images. Glaucoma – A group of eye conditions associated with increased intraocular pressure that damages the optic nerve [3]. Fundus images show changes in the optic disc such as increased cup-to-disc ratio. Myopia – A common refractive error where distant objects appear blurry [4].

Though not a disease, high myopia can lead to retinal detachment and macular degeneration, visible in fundus images. Optic Disc Cupping (ODC) – Often linked to glaucoma [5], it is characterized by the enlargement of the optic cup within the optic nerve head. It is an indicator of glaucomatous damage. Normal – Retinal images without any visible signs of pathology. Including normal class ensures the models can distinguish healthy eyes from diseased ones.

The models are trained and evaluated on a publicly available, balanced dataset of fundus images from the RFMiD repository. Evaluation is based on key performance indicators including accuracy, precision, recall, specificity, and F1-score. The key contributions of this work include:

i. Development and evaluation of a Customized CNN architecture optimized for retinal image classification.
ii. Comparative analysis with popular pretrained models (VGG16, ResNet50, DenseNet201, Efficient NetB4, ResNet152, EfficientNetB7).
iii. Implementation of a user-interactive GUI for disease prediction from fundus images.
iv. This research aims to demonstrate that deep learning models, when properly trained and validated, can serve as reliable diagnostic aids in ophthalmology.

2. LITERATURE REVIEW

A Comprehensive Review of Deep Learning Strategies in Retinal Disease Diagnosis Using Fundus Images, has been discussed in [7]. It includes, essential elements such as dataset collection, pre-processing techniques, model architectures (DenseNet, EfficientNet, and ResNet variants), and evaluation metrics used in previous studies. The review identifies persistent challenges, including handling heterogeneous image quality, balancing sensitivity and specificity, and addressing class imbalance. The review offers a comparative analysis of multiple deep learning strategies and demonstrates that robust pre-processing combined with appropriate deep neural networks can significantly improve diagnostic outcomes. It concludes that while the existing methods have achieved promising results, there is still a need for more unified frameworks that integrate feature extraction, segmentation, and classification.

The development of an AI-based diagnostic tool for retinal diseases using fundus images was discussed in [8]. Published in the *Journal of Personalized Medicine*, this study emphasized the importance of accurate and cost-effective diagnostic tools, especially for low-resource settings. The proposed deep learning model achieved an Area Under the Curve (AUC) value of 0.97 in classifying retinal conditions such as age-related macular degeneration (AMD), diabetic retinopathy (DR), glaucoma suspect (GS), and normal cases. The results suggest that such AI tools can serve as reliable, accessible, and affordable screening solutions for the early detection and differential diagnosis of retinal diseases in clinical practice. In [9], the authors introduced an ensemble-based deep learning approach for retinal disease prediction using segmented blood vessels from retinal images. The system utilized multiple deep learning models in combination to enhance accuracy and robustness. Evaluated on the DRIVE and STARE datasets, the model achieved 99.71% accuracy, 98.63% precision, 98.25% recall, and 99.22% F1-score. The extracted feature vectors were fused and classified using a softmax classifier. The study underscores the potential of ensemble-based methods in improving the reliability and effectiveness of automated retinal disease diagnosis systems. In [10] Deep Learning-Based Ocular Disease Classification in Fundus Images. Authors S. Ortiz and M. A. Goenaga Jimenez, 2023 This study develops a multiclass classification model based on a pre-trained ResNet-50 to diagnose ocular diseases from 1,096 retinal fundus images, employing advanced pre-processing and data augmentation techniques to enhance robustness. The primary challenges involve distinguishing subtle disease-specific features and addressing the limited availability of annotated data while preventing overfitting. The results demonstrate high diagnostic accuracy, affirming the efficacy of transfer learning, with future work exploring real-time imaging integration and hybrid models for improved clinical utility.

In [11] This study compares multiple CNN architectures for retinal disease detection using 3,200 fundus images, focusing on cross-validation, data augmentation, and standardized pre-processing to ensure reliable performance assessment. Challenges include class imbalance, intra-class variability, and efficient feature extraction, with DenseNet-based models excelling in vascular feature recognition while ResNet-based models demonstrated high accuracy through transfer learning. The authors conclude that model performance depends on both architecture and effective preprocessing, suggesting future research into ensemble methods and optimization techniques for

enhanced multi-disease classification. In [12], the review explores deep learning strategies for retinal disease diagnosis, covering CNN architectures, dataset collection, pre-processing techniques, and evaluation metrics. Key challenges include heterogeneous image quality, sensitivity-specificity balance, and class imbalance, with findings highlighting that strong pre-processing and well-chosen neural networks significantly enhance diagnostic accuracy. The authors emphasize the need for unified frameworks
integrating feature extraction, segmentation, and classification, while future research should focus on hybrid models, real-time processing, and improved model explainability for clinical applications.
A deep learning-based model for the classification of retinal diseases using Optical Coherence Tomography (OCT) images was proposed in [13]. The authors designed a novel Convolutional Neural Network (CNN) architecture and compared its performance against pre-trained models such as Inception V3 and VGG-16. The proposed CNN achieved an accuracy of 98.5%, while Inception V3 achieved 99.27% and the unmodified VGG-16 reached only 53%. By incorporating additional convolutional layers and regularization terms, the modified VGG-16 architecture attained an improved accuracy of 93.5%. This work highlights the effectiveness of deep learning techniques in accurately classifying different stages of retinal diseases and aiding ophthalmologists in developing patient-specific treatment plans. In [14], the study proposes an ensemble deep transfer learning strategy that trains on 38,727 high-quality fundus images and subsequently tests on 13,000 low-quality images acquired with cost-effective equipment. The ensemble leverages a combination of convolutional neural networks, integrated with transfer learning techniques, to maintain high diagnostic accuracy even when image quality is compromised. The work includes dealing with significant variations in image quality and ensuring that feature extraction remains robust across both high- and low-quality datasets. The ensemble method achieved classification accuracies ranging from approximately 79% to over 90% across various conditions (e.g., diabetic retinopathy and excavation). The study concludes that a deep transfer learning approach can effectively bridge the quality gap between training and testing data, making it especially useful in resource-constrained settings.
In summary, the surveyed works provide a robust foundation for the current thesis, which aims to further refine multi-disease retinal classification by leveraging DenseNet201, EfficientNetB4, ResNet152, and a customized CNN architecture. This literature survey not only highlights the current state of the art but also identifies key challenges and future opportunities for advancing automated retinal diagnosis using deep learning techniques.

3. METHODOLOGY

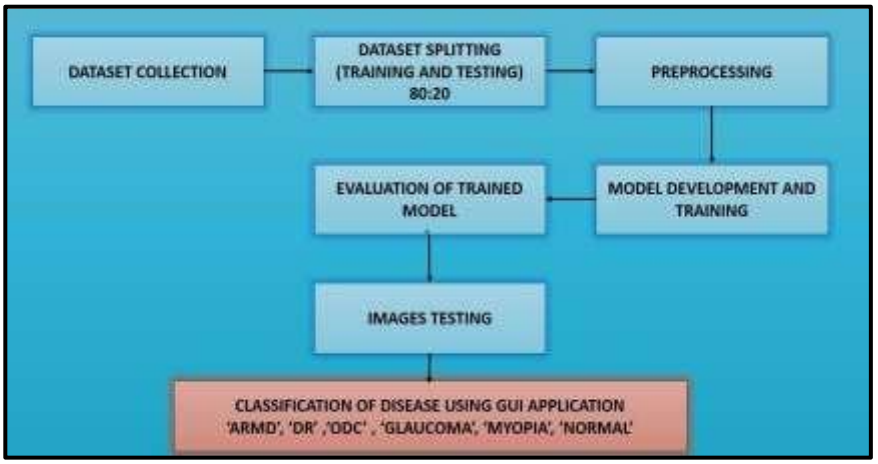

Figure 1: Block Diagram

Block diagram (Figure 1) illustrates the implemented overall architecture of the retinal disease classification system. It begins with dataset collection, followed by splitting into training and testing sets, pre processing, model training, and evaluation. The final output is the classification of six retinal conditions—ARMD, DR, ODC, Glaucoma, Myopia, and Normal—through a GUI-based application.

3.1 Dataset Collection

The dataset utilized in this study was sourced from the RFMiD (Retinal Fundus Multi-Disease Image Dataset) repository. It comprises 600 labeled retinal images, with an equal distribution across six classes: Age-Related

Macular Degeneration (ARMD), Diabetic Retinopathy (DR), Optic Disc Cupping (ODC), Glaucoma, Myopia, and Normal, containing 100 images per class.

3.2 Dataset Splitting

To facilitate robust model evaluation, the dataset was partitioned into training and test subsets in an 80:20 ratio. This split ensures a sufficient quantity of data for model training while preserving an independent set for unbiased evaluation.

3.3 Data Preprocessing

3.3.1 Image Reading: Images were read using the Python Imaging Library (PIL). The function Image.open() converts image files into NumPy arrays in RGB mode, providing a 3D representation with dimensions (height, width, channels), suitable for subsequent image processing tasks.

3.3.2 Image Resizing: All images were resized to 224×224 pixels using the OpenCV resize() function. This standard size aligns with the input requirements of several deep learning models, such as VGG16, ResNet50, and EfficientNetB7.

3.3.3 Data Augmentation: To enhance dataset variability and mitigate overfitting, the following augmentation techniques were employed: horizontal flipping, elastic transformation, optical distortion, grid distortion, and brightness adjustment.

3.4 Model Architecture

This work evaluates four convolutional neural network (CNN) architectures: VGG16, ResNet50, EfficientNetB7, DenseNet201, EfficientNetB4, ResNet152 and a custom CNN model. The models were trained and validated using the prepared dataset. These models are called transfer learning models, due to their ability to assign similar weights to new models, for training in image classification tasks [15].

3.4.1 VGG16: The VGG16 model was initialized with include_top=False to exclude the final classification layers and retain only the convolutional base. A new model was constructed by appending a Flatten layer, a Dense layer with 128 ReLU-activated units, a Dropout layer with a rate of 0.5, and a final Dense layer with 6 Softmax-activated output units, corresponding to the six target classes.

3.4.2 ResNet50: The ResNet50 model was also initialized with include_top=False and pre-trained weights from ImageNet. The base model was frozen during training to retain pre-learned features. A Sequential model was built by adding the ResNet50 base, followed by an AveragePooling2D layer, a Flatten layer, and a Dense output layer with six Softmax units. Early stopping was employed to halt training upon stagnation in validation loss.

3.4.3 EfficientNetB7: EfficientNetB7, known for its compound scaling approach, consists of 66 layers including 55 Mobile Inverted Bottleneck Convolution (MBConv) blocks, squeeze-and-excitation mechanisms, and Swish activations. The model's depth, width, and resolution are co-scaled to achieve high performance with computational efficiency

3.4.4 DenseNet201: DenseNet201, leveraging densely connected layers, consists of 201 layers that enhance feature reuse and gradient flow, utilizing bottleneck and transition layers for efficient parameter usage.

3.4.5 ResNet152: ResNet152, a deep residual network, employs 152 layers with identity shortcuts and residual connections, effectively mitigating vanishing gradients while maintaining high accuracy in image recognition tasks.

3.4.6 EfficientNetB4: EfficientNetB4 balances depth, width, and resolution using compound scaling, incorporating 48 layers with MBConv blocks and squeeze-and-excitation mechanisms to achieve strong accuracy with minimal computational cost

3.4.7 Custom CNN: A tailored 4 layer CNN architecture was designed to establish a baseline. The model comprises sequential convolutional, pooling, and dropout layers, culminating in fully connected layers. The architecture was optimized via experimentation to balance complexity and performance.

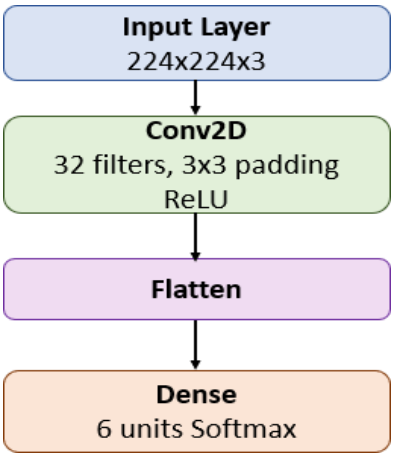

Figure 2: Customized CNN Architecture

The Customized CNN was designed specifically for retinal disease classification. The Architecture consists of four convolutional layers followed by two fully connected layers. The Activation Functions used are ReLU for intermediate layers and Softmax for the final classification layer. For regularization, Dropout and batch normalization are employed to prevent overfitting.

The network architecture is designed to balance the extraction of detailed features with computational efficiency, ensuring its suitability for complex medical imaging tasks. The key components of the architecture are as follows:

i. **Convolutional Blocks**: The network begins with one or more convolutional layers using small 3×3 kernels. These layers are tasked with detecting low-level features such as edges, textures, and micro-patterns present in retinal images. The modular design permits experimentation with different depths (ranging from two to five convolutional layers), allowing for enhanced feature extraction in deeper configurations. The ability to adjust the number of layers offers a flexible trade-off between complexity and computational cost.

ii. **Activation Functions (ReLU):** Each convolutional layer is immediately followed by a Rectified Linear Unit (ReLU) activation function. The non-linearity introduced by ReLU is vital for modeling the complex and non-linear relationships inherent in retinal data, thereby boosting the network's capacity to distinguish subtle pathological variations.

iii. **Pooling Layers**: While the document does not mandate pooling operations in every layer, max pooling is typically applied after certain convolutional blocks. Pooling reduces the spatial dimensions of the feature maps, capturing the most prominent features and contributing to translation invariance. This step is integral to managing computational loads and mitigating overfitting.

iv. **Flattening and Dense Layers:** The multidimensional data output from the convolutional blocks is flattened into a one-dimensional vector. This representation is then fed into one or more fully connected (dense) layers, culminating in an output layer equipped with a softmax activation. The softmax function converts the network's output into probability scores for each retinal disease class—enabling the final classification decision.This concise yet effective architectural design facilitates both efficient training and robust feature learning, which is particularly valuable for applications such as retinal disease classification where both local and global features play crucial roles.

v. **Training Methodology:** The architecture leverages the Keras/TensorFlow fit() method for training, processing the entire dataset over multiple epochs. This iterative training strategy allows the model to progressively adjust its weights through backpropagation, ensuring that both local and global features are optimally learned.

vi. **Hyperparameter Tuning:** Careful hyperparameter tuning was carried out using cross-validation to ensure that the model not only converges but also generalizes well to unseen data. The following are the critical hyperparameters adjusted during the tuning process:

Batch Size (128): A batch size of 128 was chosen primarily to optimize GPU utilization. By processing 128 images per iteration, the model achieves stable gradient estimates while balancing memory use and computational efficiency. Larger batch sizes can also contribute to faster convergence, although they must be chosen carefully to avoid generalization issues.

Learning Rate (0.0004): The learning rate determines the step size at each iteration as the optimizer adjusts the

weights. Setting it to 0.0004 strikes a balance between making significant progress during training and preventing the updates from overshooting the optimal values. Such a fine learning rate supports stable convergence, especially in complex models where sudden weight updates may destabilize learning.

Epochs (100, 400, and 500): The number of epochs—essentially the number of complete passes through the training dataset—varied depending on the specific architecture employed. Simpler models converged faster (around 100 epochs), while more complex architectures or those requiring finer feature extraction were trained for 400 to 500 epochs. This flexibility allowed us to tailor training duration to the model's complexity, ensuring ample learning without overfitting.

Loss Function (Categorical Cross-Entropy): Since the task involves multiclass classification (multiple types of retinal diseases), categorical cross-entropy was the appropriate choice for the loss function. This function measures the divergence between the predicted class probabilities and the actual distribution, guiding the network to improve its predictions over time.

Hyperparameter tuning through cross-validation ensured that each setting—from batch size to learning rate—was optimized for performance and generalization. It allowed for systematic testing across different partitions of data, reducing both variance and bias and, ultimately, leading to a robust CNN model well-suited for the diagnostic challenges in medical imaging.

3.5 Model Training

Model training was conducted using the fit() function in Keras. Models were trained over 100 epochs with a batch size of 32,64. The Adam optimizer was used, and learning rates were fine-tuned for optimal convergence.

3.6 Model Evaluation

The trained models were evaluated using several performance metrics:

3.6.1 Confusion Matrix: The confusion matrix was computed using the confusion_matrix() function from Scikit- learn. A heatmap representation was plotted using Seaborn to visualize the distribution of predicted versus actual labels.

Accuracy: Accuracy measures the ratio of correctly predicted instances to the total instances.

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \quad \text{---------} (1)$$

Precision: Precision quantifies the correctness of positive predictions and was computed using precision_score() with weighted averaging.

$$Precision = \frac{TP}{TP + FP} \quad \text{- - - - - - - -} (2)$$

Recall: Recall assesses the model's ability to identify all relevant instances.

$$Recall = \frac{TP}{TP + FN} \quad \text{- - - - - - - -} (3)$$

F1-Score: F1-score provides a harmonic mean of precision and recall, computed using f1_score() with a weighted average.

$$F1\ Score = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad \text{---------} (4)$$

Specificity: Specificity evaluates the model's ability to correctly identify negative instances.

$$Specificity = \frac{TN}{TN + FP} \quad \text{----------} (5)$$

3.7 Testing and Algorithm Comparison

After training, each model's performance was evaluated using the test set. Models were compared based on the

aforementioned metrics to identify the most effective architecture for multi-class retinal disease classification**.**

4. RESULTS AND DISCUSSION

4.1 CNN Model Implemented

This section details the performance evaluation of four deep learning models VGG16, ResNet50, EfficientNetB7, DenseNet201, EfficientNetB4, ResNet152 and a Customized CNN the classification of five types of retinal diseases with normal: ARMD, DR, ODC, GLAUCOMA, MYOPIA, and NORMAL. All models were trained and tested on the RFMiD dataset, which includes 600 fundus images equally distributed across the six categories. The dataset was split in an 80:20 ratio for training and testing, respectively. The evaluation metrics included training and validation accuracy, loss, confusion matrix, sensitivity, specificity, precision, and F1-score to assess the classification performance of each model.

4.1.1 VGG16

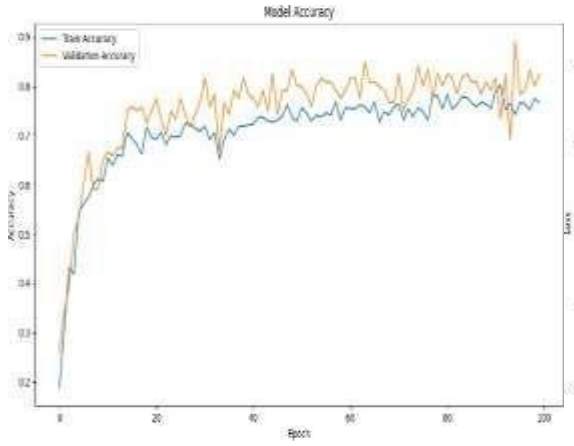

Figure 3: VGG16 Accuracy Graph

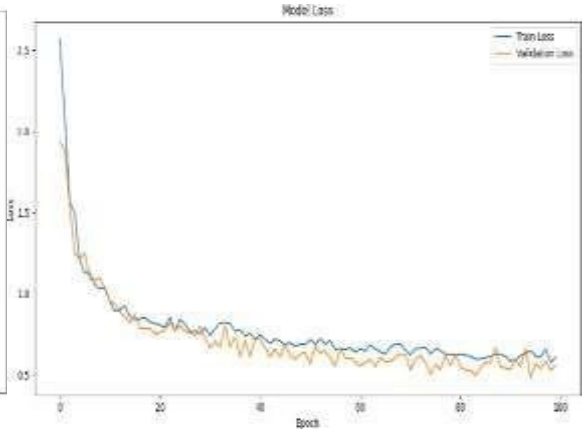

Figure 4: VGG16 Loss Graph

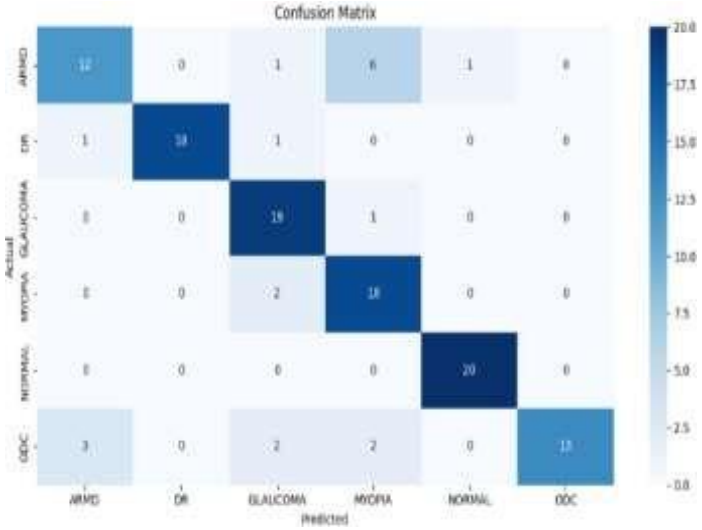

Figure 5: VGG16 Confusion Matrix

The VGG16 model demonstrated a reliable learning pattern during training at 100 epochs, it achieved a training accuracy of 76% and validation accuracy of 82% (Figure 3)**.** The accuracy and loss curves (Figure 4), indicated steady convergence over training epochs. The confusion matrix showed (Figure 5) that VGG16 performed well in classifying DR, GLAUCOMA, MYOPIA and NORMAL images, with a few misclassifications observed in ARMD and ODC categories. The model struggled with complex feature variations due to its relatively shallow architecture and lack of residual connections.

International Journal of Environmental Sciences
ISSN: 2229-7359
Vol. 11 No. 20s, 2025
https://www.theaspd.com/ijes.php

4.1.2 ResNet50

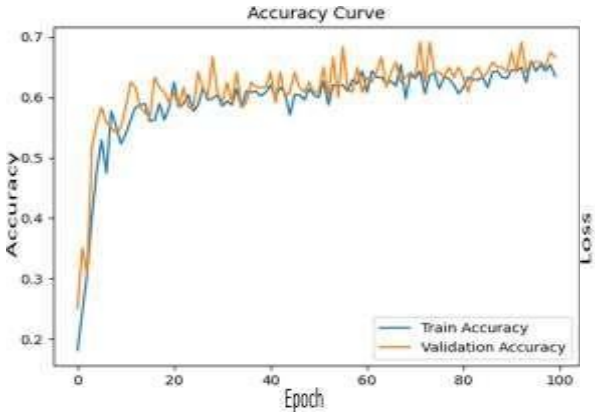

Figure 6: ResNet50 Accuracy Graph

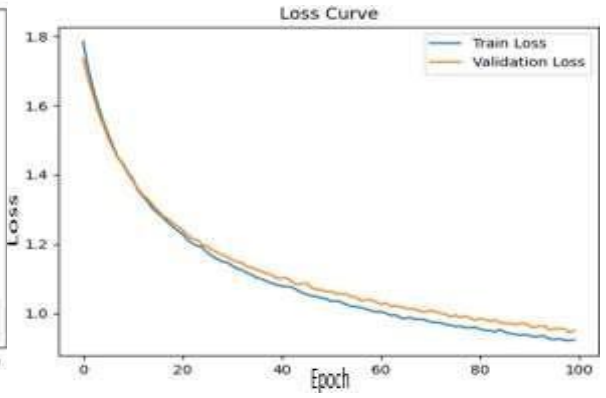

Figure 7: ResNet50 Loss Graph

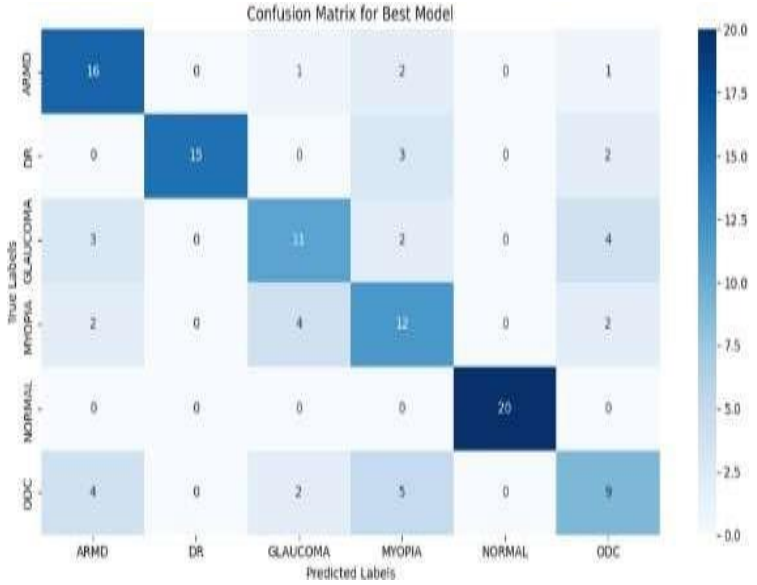

Figure 8: ResNet50 Confusion Matrix

ResNet50, with its residual learning mechanism showed improved performance over VGG16, achieving a maximum training accuracy of 70% and validation accuracy of 69% at 75 epochs (Figure 6), with smoother loss reduction (Figure 7). The confusion matrix reflected (Figure 8) enhanced classification performance, especially in distinguishing ARMD. ResNet50's deeper architecture enabled it to extract more representative features from fundus images, leading to better overall generalization.

4.1.3 EfficientNetB7

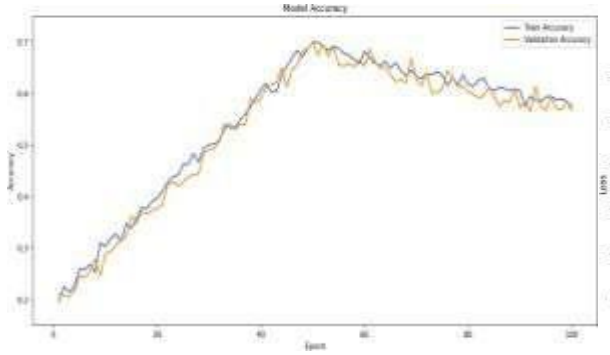

Figure 9: EfficientNetB7 Accuracy Graph

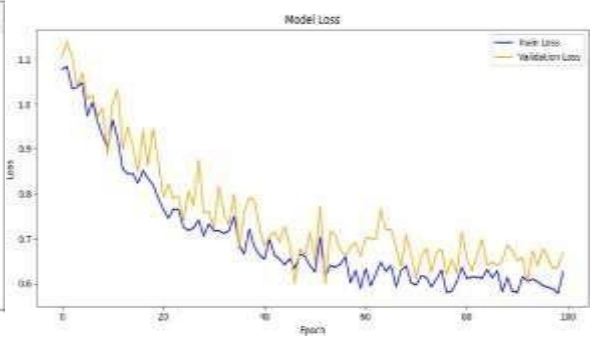

Figure 10: EfficientNetB7 Loss Graph

International Journal of Environmental Sciences
ISSN: 2229-7359
Vol. 11 No. 20s, 2025
https://www.theaspd.com/ijes.php

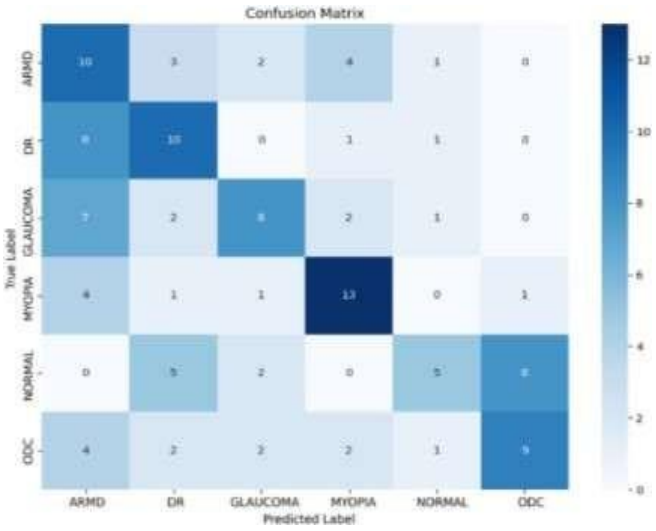

Figure 11: EfficientNetB7 Confusion Matrix

EfficientNetB7, known for its compound scaling efficiency, reached a training accuracy of 70% and validation accuracy of 69% at 50 epochs (Figure 9). The training and validation graphs revealed minimal overfitting and efficient learning. The confusion matrix indicated (Figure 11) accurate classification across all six classes, with the lowest rate of misclassification among the pretrained models. This model also showed consistent validation accuracy and low validation loss (Figure 10), affirming its superior performance.

4.1.4 Efficient Net-B4

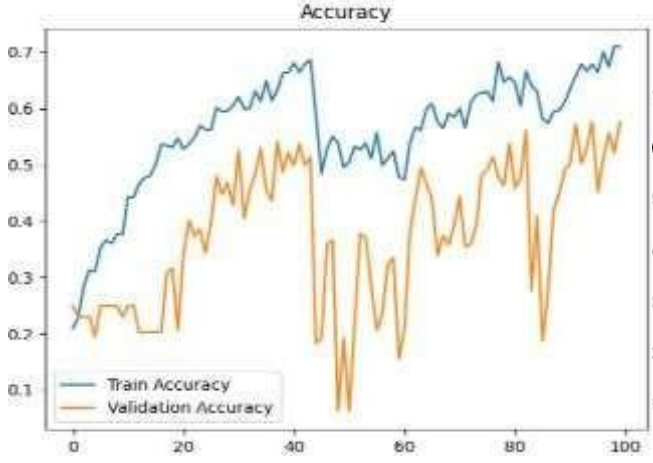

Figure 12: EfficientNetB4 Accuracy Graph

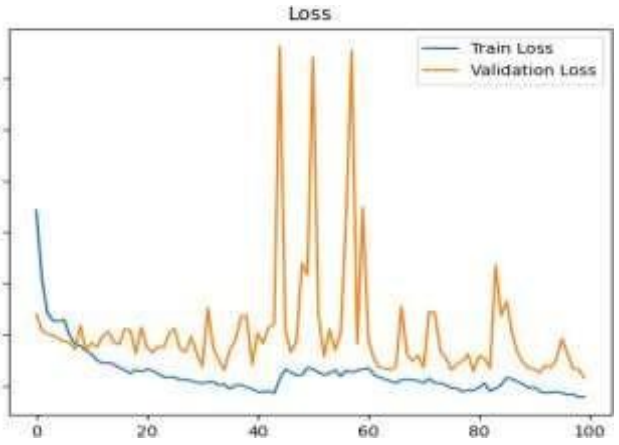

Figure 13: EfficientNetB4 loss Graph

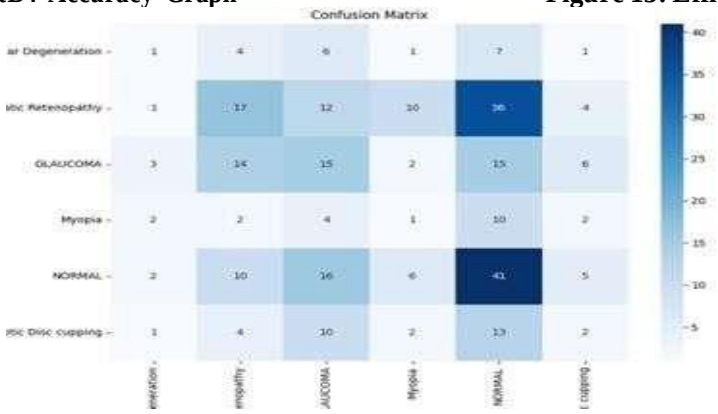

Figure 14: EfficientNetB4 Confusion Matrix

The EfficientNet-B4 model exhibited a consistent learning trajectory over 100 training epochs, achieving an overall accuracy of 73% (Figure 12). The accuracy and loss curves (figure 13), revealed steady convergence throughout the training process. Analysis of the confusion matrix indicated (Figure 14) that the model effectively classified images belonging to the DR, Glaucoma, Myopia, and Normal categories, although some misclassifications were noted in the ARMD and ODC classes. These results suggest that while EfficientNet-B4 is competent for certain retinal conditions, its relatively shallow architecture and absence of residual connections may limit its ability to capture complex feature variations in more challenging categories.

4.1.5 ResNet152

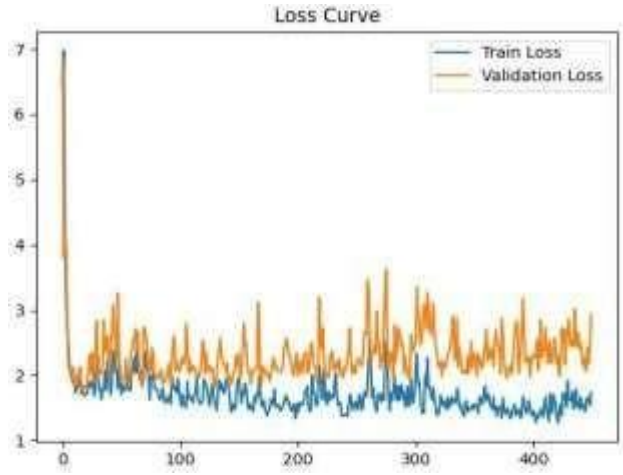

Figure 15: ResNet152 Accuracy Graph

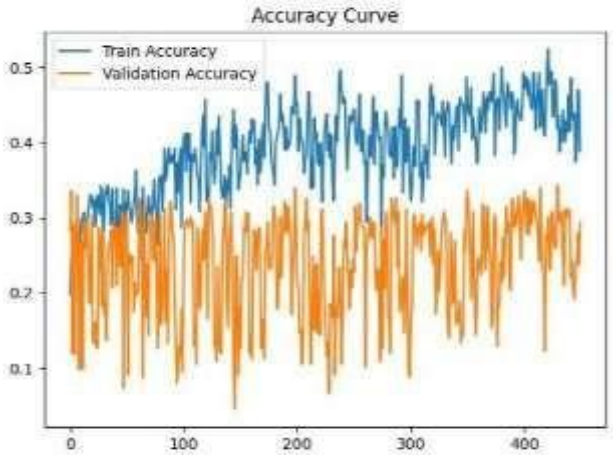

Figure 16: ResNet152 loss Graph

Figure 17: ResNet152 Confusion Matrix

ResNet152, leveraging its residual learning mechanism, achieved a peak accuracy of 50% after 400 epochs (Fig 15), which did not improve upon the performance observed with EfficientNet-B4. Despite this, its training process showed a smoother loss reduction (Fig 16), indicating stable convergence. Furthermore, analysis of the confusion matrix revealed (Fig 17) that ResNet152 distinguishing ARMD cases, suggesting that its deeper architecture enabled the extraction of more representative and discriminative features from the fundus images.

International Journal of Environmental Sciences
ISSN: 2229-7359
Vol. 11 No. 20s, 2025
https://www.theaspd.com/ijes.php

4.1.6 DenseNet201

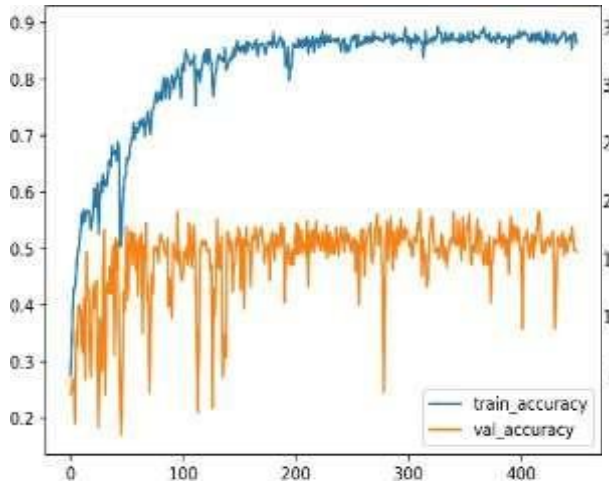

Figure 18: DenseNet201 Accuracy Graph

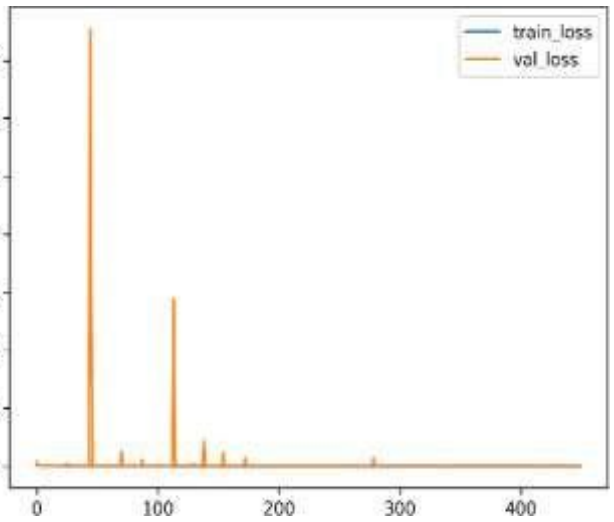

Figure 19: DenseNet201 loss Graph

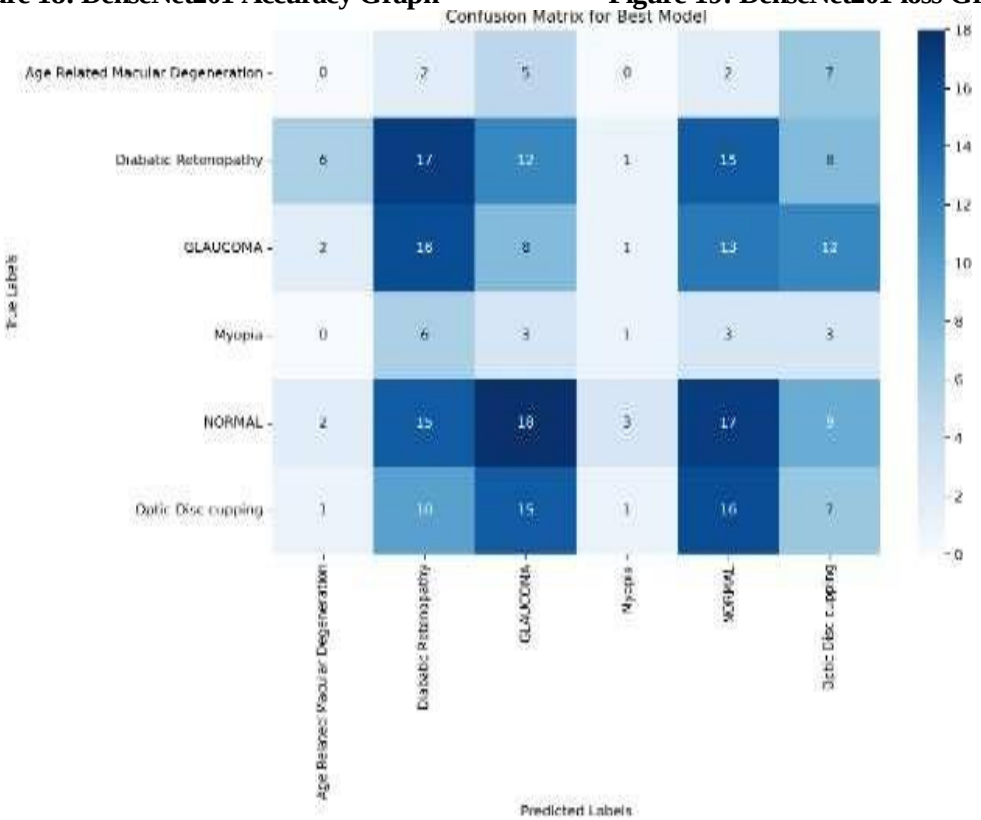

Figure 20: DenseNet201 Confusion Matrix

DenseNet201, leveraging its densely connected architecture, demonstrated promising performance by achieving an overall training accuracy of 87% and validation accuracy of 68% over 400 epochs (Figure 18). Its intrinsic design— characterized by efficient gradient propagation and feature reuse—facilitated rapid convergence and enabled the network to extract intricate, hierarchical features from fundus images. Analysis of the confusion matrix revealed (Figure 20) that DenseNet201 excelled in accurately classifying conditions such as diabetic retinopathy, glaucoma, and normal retinal images, although it showed some misclassifications in distinguishing ARMD and ODC cases. Overall, while DenseNet201's architecture effectively addresses complex feature extraction, its performance is still somewhat affected by challenges related to class imbalance and subtle inter-class variations.

International Journal of Environmental Sciences
ISSN: 2229-7359
Vol. 11 No. 20s, 2025
https://www.theaspd.com/ijes.php

4.1.7 Customized CNN

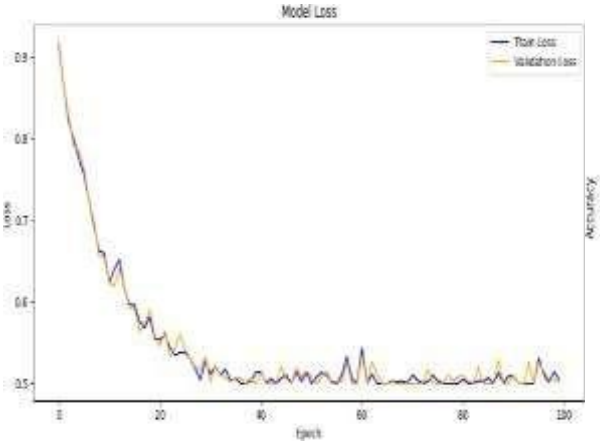

Figure 21: Customized CNN Accuracy Graph

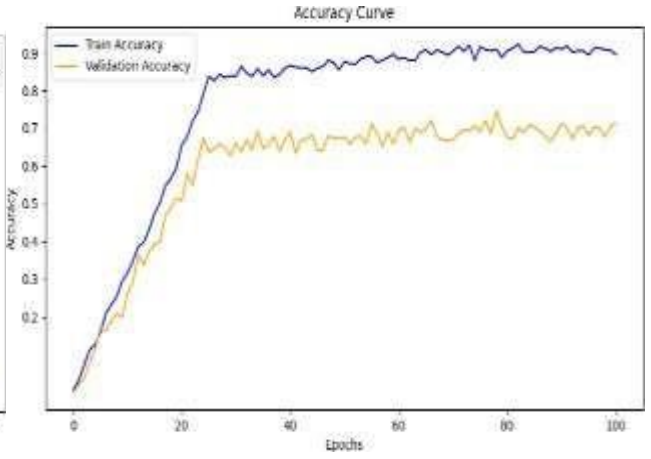

Figure 22: Customized CNN Loss Graph

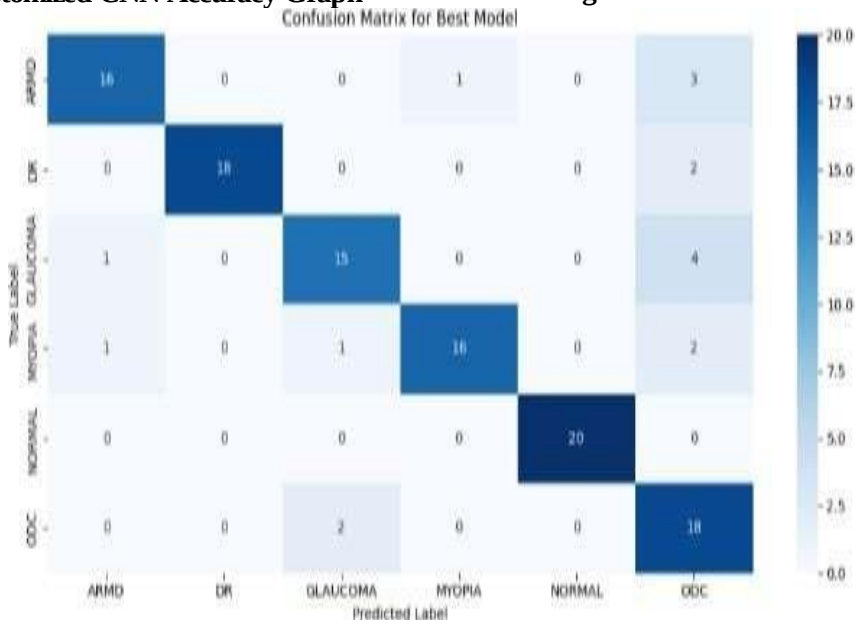

Figure 23: Customized Confusion Matrix

A custom CNN was built and tested with varying depths. The best performance was achieved using a four-layer architecture The Customized CNN outperformed all other models. At 100 epochs, it achieved a training accuracy of 90.8% and validation accuracy of 70% (Figure 21), with consistent loss values around 0.50 (Figure 22). Analysis of the confusion matrix revealed (Figure 23) that excelled in accurately classifying conditions After tuning the number of layers and epochs, the model achieved competitive results. The custom model had fewer parameters and a lighter architecture, making it suitable for real-time applications and deployment in edge devices.

4.2 Comparative Summary

Epochs	VGG16	ResNet50	EfficientNet B7	EfficientNet B4	ResNet 152	DenseNet 201	Custom CNN
25	0.70	0.58	0.47	0.57	0.34	0.40	0.83
50	0.72	0.62	0.70	0.48	0.29	0.42	0.87
75	0.74	0.70	0.63	0.61	0.36	0.45	0.907
100	0.76	0.62	0.59	0.73	0.39	0.50	0.908

Table 1: Comparision of Accuracy for various models for 25,50,75 and 100 Epochs

Epochs	VGG16	ResNet50	EfficientNet B7	EfficientNet B4	ResNet 152	DenseNet 201	Custom CNN
25	0.84	1.21	0.74	1.09	1.21	1.50	0.53
50	0.66	1.02	0.67	0.66	1.02	0.45	0.50
75	0.64	0.64	0.64	0.44	0.84	0.64	0.50
100	0.11	0.91	0.61	0.11	0.38	0.41	0.50

Table 2: Comparision of Loss for various models for 25,50,75 and 100 Epochs

Between 25 and 100 epochs, the Custom CNN's accuracy increases markedly from 83 percent to 90.8 percent, thereby outperforming all pretrained architectures—VGG16 (70–76 percent) and EfficientNetB4 (57–73 percent) among them—which exhibit more modest gains (Table 1). Moreover, the Custom CNN maintains the lowest training loss (Table 2) at every checkpoint, whereas deeper models such as ResNet152 and EfficientNetB7 record consistently higher loss, reflecting slower convergence and a possible mismatch between model capacity and data volume.

Epochs	VGG16	ResNet50	EfficientNet B7	EfficientNetB4	ResNet1 52	DenseNet 201	Custom CNN
25	0.73	0.58	0.44	0.59	0.34	0.39	0.65
50	0.80	0.62	0.69	0.50	0.30	0.42	0.67
75	0.84	0.69	0.61	0.62	0.31	0.47	0.69
100	0.82	0.67	0.57	0.70	0.32	0.48	0.70

Table 3: Comparision of Val_ Accuracy for various models for 25,50,75 and 100 Epoch

Epochs	VGG16	ResNet50	EfficientNet B7	EfficientNetB4	ResNet1 52	DenseNet 201	Custom CNN
25	0.80	1.19	0.80	1.08	1.32	1.48	0.54
50	0.63	1.02	0.73	0.45	1.05	0.54	0.51
75	0.60	0.99	0.71	0.35	0.08	0.46	0.51
100	0.55	0.94	0.70	0.61	0.12	0.14	0.50

Table 4: Comparision of Val_Loss for various models for 25,50,75 and 100 Epoch

Across 25 to 100 epochs on unseen data, VGG16 achieves its highest validation accuracy of 84 percent at 75 epochs, while the Custom CNN stabilizes around 70 percent and architectures such as ResNet50 and EfficientNetB7 record even lower accuracies, reflecting divergent generalization and overfitting behaviors. Simultaneously, validation loss trajectories reveal that VGG16 and ResNet152 attain the lowest errors by mid-training, whereas the Custom CNN's consistently moderate loss (Tables 3 & 4) suggests reliable stability with potential for further generalization enhancement.

4.3 Comparision of Evaluation Metrics for various models

Accuracy							
	VGG16	**ResNet50**	**EfficientNet B7**	**EfficientNet B4**	**ResNet1 52**	**DenseNet 201**	**Custom CNN**
AMRD	0.80	0.83	0.75	0.90	0.88	0.89	0.95
DR	0.84	0.83	0.65	0.66	0.72	0.65	0.98
GLAUCOMA	0.85	0.74	0.74	0.69	0.81	0.62	0.93
MYOPIA	0.82	0.68	0.76	0.86	0.60	0.92	0.96
NORMAL	0.89	0.61	0.70	0.58	0.80	0.63	1.00
OCD	0.83	0.83	0.71	0.83	0.86	0.68	0.89

Table 5: Accuracy

The Custom CNN dominates per-class accuracy (0.89–1.00), underlining its balanced performance across AMRD, DR, Glaucoma, Myopia, Normal, and ODC. Pretrained models show greater variance, with EfficientNetB7 and ResNet152 particularly struggling on DR and ODC.

Sensitivity							
	VGG 16	**ResNet 50**	**EfficientNet B7**	**EfficientNet B4**	**ResNet152**	**DenseNet 201**	**Custo m CNN**
AMRD	0.55	0.00	0.00	0.05	0.00	0.00	0.80
DR	0.35	0.00	0.12	0.21	0.01	0.29	0.90
GLAUCOMA	0.40	0.30	0.30	0.27	0.07	0.15	0.75
MYOPIA	0.80	0.40	0.25	0.05	0.80	0.06	0.80
NORMAL	0.55	0.90	0.90	0.51	0.95	0.27	1.00
OCD	0.45	0.00	0.00	0.83	0.03	0.14	0.90

Table 6: Sensitivity

Custom CNN achieves high true-positive rates (0.75–1.00), ensuring reliable detection of all six conditions, including perfect recall for Normal cases. In contrast, many pretrained nets record near-zero sensitivity in several classes, indicating frequent misses on positive samples.

Specificity							
	VGG 16	**ResNet 50**	**EfficientNet B7**	**EfficientNet B4**	**ResNet 152**	**DenseNet 201**	**Custom CNN**
AMRD	0.85	1.00	0.90	0.97	0.95	0.95	0.98
DR	0.94	1.00	0.74	0.84	0.99	0.75	1.00
GLAUCOM A	0.94	0.83	0.83	0.79	0.99	0.74	0.97
MYOPIA	0.82	0.74	0.86	0.92	0.59	0.98	0.99
NORMAL	0.96	0.55	0.55	0.61	0.75	0.75	1.00
OCD	0.91	1.00	0.12	0.93	0.97	0.81	0.89

Table 7: Specificity

With specificity scores above 0.97 across most categories, the Custom CNN excels at avoiding false alarms, especially in DR and Normal classes. While VGG16 also performs well overall, models like EfficientNetB7 and ResNet152 exhibit dips (e.g., NORMAL at 0.55), revealing weaknesses in true-negative identification.

Precision							
	VGG 16	**ResNet 50**	**EfficientNet B7**	**EfficientNet B4**	**ResNet 152**	**DenseNet 201**	**Custom CNN**
AMRD	0.42	0.00	0.00	0.10	0.00	0.00	0.89
DR	0.54	0.00	0.09	0.33	0.33	0.29	1.00
GLAUCOMA	0.57	0.26	0.26	0.24	0.67	0.13	0.83
MYOPIA	0.47	0.24	0.27	0.05	0.13	0.14	0.94
NORMAL	0.73	0.29	0.29	0.34	0.59	0.26	1.00
OCD	0.50	0.00	0.12	0.10	0.11	0.15	0.62

Table 8: Precision

Precision metrics underscore the Custom CNN's correctness of positive predictions—achieving 1.00 for DR and Normal—minimizing false positives. Other architectures frequently record zero or low precision in diseases such as AMRD and OCD, highlighting unbalanced predictive behavior.

F1-score							
	VGG 16	**ResNet 50**	**EfficientNet B7**	**EfficientNet B4**	**ResNet 152**	**DenseNet 201**	**Custom CNN**
AMRD	0.48	0.00	0.00	0.07	0.00	0.00	0.84
DR	0.42	0.00	0.11	0.26	0.02	0.27	0.95
GLAUCOMA	0.47	0.28	0.28	0.25	0.13	0.14	0.79
MYOPIA	0.59	0.30	0.30	0.05	0.22	0.09	0.86
NORMAL	0.63	0.43	0.43	0.41	0.73	0.26	1.00
OCD	0.47	0.00	0.12	0.08	0.05	0.15	0.73

Table 9: F1-score

International Journal of Environmental Sciences
ISSN: 2229-7359
Vol. 11 No. 20s, 2025
https://www.theaspd.com/ijes.php

By combining precision and recall, the Custom CNN secures the highest F1-scores (0.73–1.00), reflecting both accurate and consistent classification. Pretrained models yield low F1 values in multiple classes, demonstrating uneven trade-offs between sensitivity and precision.

4.4 GUI Result

Figure 24: GUI of the Retinal Disease Classifier

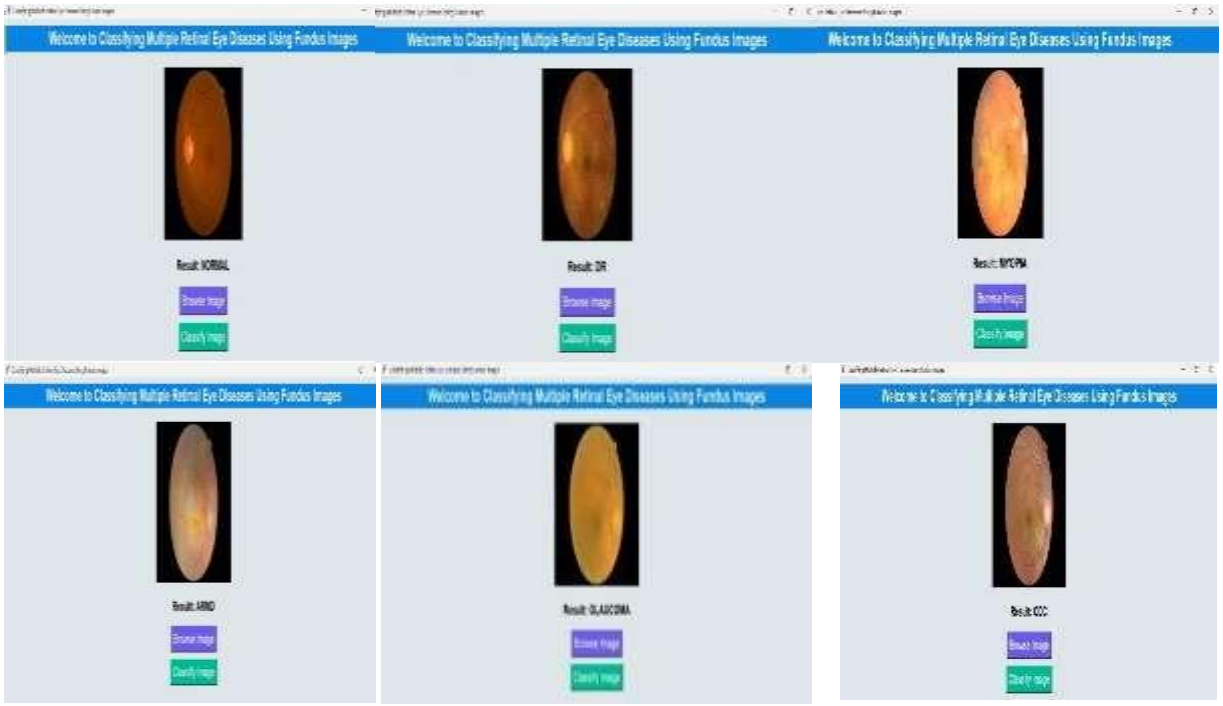

Fig 25: Classifier showing the predictions on an image

The GUI Image Classification of Retinal Diseases is an interactive system that integrates a Customized CNN within a user-friendly interface, enabling clinicians to upload and analyze retinal images for accurate disease prediction. A further breakdown by class illustrates that the Customized CNN consistently outperforms across all metrics— Accuracy, Sensitivity, Specificity, Precision, and F1-score—as showcased by the classifier's GUI display (Figure 24) and its prediction output (Figure 25).

5. CONCLUSION AND FUTURE SCOPE

The detection and classification of eye abnormalities is important in clinical treatments. With increase in population there is a pressing need for an automated system that detects and classifies eye diseases with higher efficiency. The eye diseases when not detected at an early stage will lead to vision loss. These diseases do not show symptoms at an early stage. Therefore, there is a need to detect these diseases at early stages to prevent vision loss. The proposed model detects and classifies three major eye abnormalities which includes ARMD, DR, and ODC. The model provides training and testing accuracy of 95.24% and 81.68% respectively. Future scope may include the use of sophisticated classification techniques with a large dataset. For quick and dependable solutions, a similar strategy might be used to other medical imaging issues. Better augmentation methods can be used to increase the number of images. One of the most powerful techniques used is Generative Adversarial Networks (GAN). They consist of a Generator, which creates fake data, and a Discriminator, which distinguishes

between real and fake data. Through adversarial training, the Generator learns to produce increasingly realistic data, making it valuable for augmenting small or imbalanced datasets. This technique is especially beneficial in fields where data collection is expensive or time-consuming, such as medical imaging. U-Net can also be used for classification. Its unique structure consists of a symmetric U-shaped design with a contracting path (encoder) and an expansive path (decoder). The encoder captures context and features through a series of convolutional and max-pooling layers, while the decoder reconstructs the spatial dimensions using up-sampling and convolutional layers.

REFERENCES

[1] P. Chakraborty, and C. Tharini, "Pneumonia and eye disease detection using convolutional neural networks." Engineering, Technology & Applied Science Research, June 2020, vol 10 (3), pp. 5769–5774.

[2] Peng Y, Dharssi S, Chen Q, Keenan TD, Agro´n E, Wong WT, Chew EY, Zhiyong Lu (2019) DeepSeeNet: "a deep learning model for automated classification of patient-based age-related macular degeneration severity from color fundus photographs". Ophthalmology 126(4):565–575.

[3] Sarki R, Ahmed K, Zhang Y (2020) "Early detection of diabetic eye disease through deep learning using fundus images". EAI Endorsed Trans Perv Health Technology 6(22).

[4] Mahum R, Rehman SU, Okon OD, Alabrah A, Meraj T, Rauf HT (2021) "A novel hybrid approach based on deep CNN to detect glaucoma using fundus imaging". Electronics 11(1):26.

[5] Russo, Andrea, Boldini, Alessandro, Romano, Davide, Mazza, Giuseppina, Bignotti, Stefano, Morescalchi,Francesco, Semeraro, "Myopia: Mechanisms and Strategies to Slow Down Its Progression", Journal of Ophthalmology, volume 2022, Issue 1, 1004977, June 2022.

[6] A. A. Lumbantoruan, A. Bustamam and P. Anki, "Retinal Disease for Clasification -Multilabel with Applying Convolutional Neural Networks Based Support Vector Machine and DenseNet," 2021 4th International Seminar on Research of Information Technology and Intelligent Systems (ISRITI), Yogyakarta, Indonesia, 2021, pp. 475-479, doi: 10.1109/ISRITI54043.2021.9702861.

[7] S. Pachade, P. Porwal, D. Thulkar, M. Kokare, G. Deshmukh, V. Sahasrabuddhe, L. Giancardo, G. Quellec, and F. Meriaudeau, 2021. Retinal fundus multi-disease image dataset (rfmid): A dataset for multi-disease detection research. *Data*, 6 (2), p. 14.

[8] B. Goutam, M. F. Hashmi, Z. W. Geem and N. D. Bokde, "A Comprehensive Review of Deep Learning Strategies in Retinal Disease Diagnosis Using Fundus Images," in IEEE Access, vol. 10, pp. 57796-57823, 2022.

[9] Kim, K.M., Heo, T.Y., Kim, A., Kim, J., Han, K.J., Yun, J. and Min, J.K., 2021. "Development of a fundus image-based deep learning diagnostic tool for various retinal diseases". Journal of Personalized Medicine, 11(5), p.321.

[10] Kumar, K.S. and Singh, N.P., 2023. "Retinal disease prediction through blood vessel segmentation and classification using ensemble-based deep learning approaches". Neural Computing and Applications, pp.1-17.

[11] Ling-Ping Cen 1,6, Jie Ji2,3,4,6, Jian-Wei Lin1. "Automatic detection of 39 fundus diseases and conditions in retinal photographs using deep neural networks". ARTICLE NATURE COMMUNICATIONS | (2021).

[12] S. Nawaz, M., Rashid, J., Mahum, R., Masood, M., Mehmood, A., Ali, F., Kim, J., Kwon, H.Y. and Hussain, A., "Detection of diabetic eye disease from retinal images using a deep learning based CenterNet model". Sensors, 21(16), p.5283. 2021.

[13] S. Ortiz and M. A. Goenaga Jimenez, "Deep Learning-Based Ocular Disease Classification in Fundus Images," 2023 IEEE Colombian Caribbean Conference (C3), Barranquilla, Colombia, 2023, pp. 1-6.

[14] Berrimi, M. and Moussaoui, A., 2020, October. "Deep learning for identifying and classifying retinal diseases". In 2020 2nd International Conference on computer and information sciences (ICCIS) (pp. 1- 6). IEEE.

[15] G. D. A. Aranha, R. A. S. Fernandes and P. H. A. Morales, "Deep Transfer Learning Strategy to Diagnose Eye-Related Conditions and Diseases: An Approach Based on Low Quality Fundus Images," in IEEE Access, vol. 11, pp. 37403-37411, 2023.

[16] V. Singh, M.K. Gourisaria, H. G. M., and V. Singh, 2021. "Mycobacterium tuberculosis detection using CNN ranking approach.", In *Advanced Computational Paradigms and Hybrid Intelligent Computing: Proceedings of ICACCP 2021* (pp. 583–596). Singapore : Springer Singapore.

[17] H. D. Verma, M. K. Gourisaria, S. Ghosh and B. K. Dewangan, "Comparative Analysis of CNN Models for Retinal Disease Detection," 2023 International Conference on Network, Multimedia and Information Technology (NMITCON), Bengaluru, India, 2023, pp. 1-6.

[18] Sengar, Neha & Joshi, Rakesh Chandra & Dutta, Malay Kishore & Burget, Radim. (2023). EyeDeep-Net: a multi-class diagnosis of retinal diseases using deep neural network. Neural Computing and Applications. 35. 10.1007/s00521-

023-08249-x.

[19] Malik, S., Kanwal, N., Asghar, M.N., Sadiq, M.A.A., Karamat, I. and Fleury, M., "Data driven approach for eye disease classification with machine learning". Applied Sciences, 9(14), p.2789. 2019.

[20] Zhang Z, Ji Z, Chen Q, Yuan S, Fan W (2021) "Joint optimization of CycleGAN and CNN classifier for detection and localization of retinal pathologies on color fundus photographs". IEEE J Biomed Health Inform 26(1):115–126.

[21] Sheet SS, Mohammed TST, Asari MA, Hitam WHW, Sia JSY (2022) "Retinal disease identification using upgraded CLAHE filter and transfer convolution neural network". ICT Express 8(1):142–150.

[22] Tayal A, Gupta J, Solanki A, Bisht K, Nayyar A, Masud M (2021) "DL-CNN-based approach with image processing techniques for diagnosis of retinal diseases". Multimedia Syst. pp.1–22 25.

[23] Subramanian M, Kumar MS, Sathishkumar VE, Prabhu J, Karthick A, Ganesh SS, Meem MA (2022) "Diagnosis of retinal diseases based on Bayesian optimization deep learning network using optical coherence tomography images". ComputIntell Neu rosci. Vol. 2022.

[24] Saroj SK, Kumar R, Singh NP (2020) Frechet "PDF based matched filter approach for retinal blood vessels segmentation". Computer Methods Programs Biomed 194:105490.

[25] Sahu S, Singh AK, Ghrera S P, Elhoseny M, "An approach for de-noising and contrast enhancement of retinal fundus image using CLAHE". Optics and Laser Technology, February 2019, pp: 87 –98.