RetinaVision: An Efficient Deep Learning Framework for Automated Multi-Disease Retinal Image Classification Using MobileNetV2

Gourikrishna O S
Department of Computer Science, Prajyoti Niketan College, Pudukad, University of Calicut, Kerala, India

***Abstract*—Early identification of retinal diseases is crucial in order to prevent any vision-related impairment and also decrease the chances of permanent blindness. Nevertheless, manual analysis of retinal fundus images is both time-consuming and difficult for inexperienced ophthalmologists; thus, performing large-scale screening is problematic. In this paper, RetinaVision is proposed as an intelligent deep learning solution for the automation of retinal disease classification based on retinal fundus images. The proposed system uses the MobileNetV2 architecture with transfer learning for the fast and effective classification of several kinds of retinal diseases from the RFMiD (Retinal Fundus Multi-Disease Image Dataset). Input images are pre-processed, namely, they are resized, normalized, and also augmented in order to make the model more robust and generalizable. Training of the model is performed by means of the Adam optimization algorithm using the Binary Cross-Entropy loss function for multi-label classification. Experimental evaluation shows that the suggested approach is capable of achieving high-classification performance with an overall accuracy equal to 97.09% and a precision, recall, and F1-score of 80.73%, 41.33%, and 54.67%, respectively, on the test dataset. Thus, RetinaVision provides an affordable, fast, and accurate solution for helping ophthalmologists with diagnosing retinal diseases.**

***Index Terms*—Artificial Intelligence, Deep Learning, MobileNetV2, Multi-Label Classification, RFMiD Dataset, Retinal Disease Detection, RetinaVision, Transfer Learning.**

I. INTRODUCTION

Retinal diseases are among the leading causes of visual impairment and blindness worldwide, affecting millions of people each year. Conditions such as Diabetic Retinopathy (DR), Macular Hole (MH), Optic Disc Cupping (ODC), Drusen (DN), and Tessellated Fundus (TSLN) can progressively damage the retina if left undetected. Early diagnosis is essential for preventing irreversible vision loss and improving treatment outcomes. However, manual examination of retinal fundus images is a time-consuming process that requires experienced ophthalmologists, making large-scale screening difficult, particularly in resource-limited healthcare settings.

The increasing prevalence of diabetes and age-related eye disorders has created a growing demand for automated retinal disease screening systems. In many developing regions, limited access to specialized ophthalmic care delays diagnosis and treatment, increasing the risk of permanent vision impairment. Computer-aided diagnosis based on retinal fundus image analysis offers a promising solution by enabling rapid, accurate, and cost-effective screening, thereby supporting clinicians in making timely diagnostic decisions.

Recent advances in deep learning have significantly improved the performance of automated medical image analysis. Convolutional Neural Networks (CNNs) have demonstrated remarkable capability in extracting discriminative visual features from retinal fundus images for disease classification. Among these architectures, MobileNetV2 provides an excellent balance between classification accuracy and computational efficiency through its lightweight design and depthwise separable convolutions, making it well suited for deployment in real-time and resource-constrained clinical environments.

Although several deep learning models have been proposed for retinal disease detection, many existing approaches rely on computationally intensive architectures or focus on a limited number of retinal

disorders, reducing their practicality for real-world screening applications. Furthermore, lightweight models capable of accurately identifying multiple retinal diseases remain relatively underexplored. To address these challenges, this paper presents RetinaVision, an AI-powered retinal disease detection framework based on MobileNetV2 and transfer learning. The proposed system utilizes the RFMiD (Retinal Fundus Multi-Disease Image Dataset) to perform automated multi-label retinal disease classification after image preprocessing and augmentation.

The main contributions of this work are: (1) the development of a lightweight deep learning framework for automated multi-label retinal disease detection using MobileNetV2 and transfer learning; (2) the implementation of an effective preprocessing and data augmentation pipeline to enhance model generalization; and (3) a comprehensive experimental evaluation demonstrating the effectiveness of the proposed RetinaVision framework using accuracy, precision, recall, F1-score, and confusion matrix analysis for retinal disease classification.

II. RELATED WORK

Gulshan et al. [8] introduced one of the first large-scale deep learning systems for automated diabetic retinopathy detection using retinal fundus photographs, demonstrating performance comparable to certified ophthalmologists. Ting et al. [9] extended this work by developing a deep learning framework capable of detecting diabetic retinopathy and other retinal diseases across multiethnic populations, highlighting the potential of AI-assisted ophthalmic screening.

Public retinal image datasets have played a significant role in advancing automated retinal disease classification. Decencière et al. [10] introduced the MESSIDOR database for diabetic retinopathy assessment, while Porwal et al. [11] presented the IDRiD dataset containing pixel-level annotations for diabetic retinopathy lesions. Pachade et al. [12] later released the RFMiD dataset, providing multi-disease retinal fundus images for comprehensive retinal disease classification.

Transfer learning has become the dominant approach for retinal image analysis due to the limited availability of annotated medical datasets. Pratt et al. [13] employed a Convolutional Neural Network for diabetic retinopathy classification and achieved high diagnostic accuracy using transfer learning. Gargeya and Leng [14] proposed a deep CNN-based framework for automated diabetic retinopathy detection, demonstrating reliable performance for large-scale screening. Li et al. [15] evaluated several pre-trained CNN architectures, including ResNet and DenseNet, and reported improved classification performance through transfer learning and image augmentation.

Recent studies have focused on lightweight and computationally efficient deep learning models for retinal disease detection. Howard et al. [16] proposed MobileNetV2, which utilizes inverted residual blocks and depthwise separable convolutions to achieve high accuracy with significantly fewer parameters than conventional CNNs. Tan and Le [17] introduced EfficientNet, demonstrating that compound model scaling substantially improves image classification performance while maintaining computational efficiency, making it suitable for various medical image analysis tasks.

III. ARCHITECTURE OF NAIL DISEASE CLASSIFICATION SYSTEM

The proposed RetinaVision algorithm aims at the automatic detection and classification of retinal diseases through the use of the fundus images with the help of the MobileNetV2 deep learning algorithm, which uses transfer learning. The procedure starts with the capturing of retinal fundus images from the RFMiD (Retinal Fundus Multi-Disease Image Dataset). Subsequently, image processing is performed on the retinal images through the process of resizing, normalization, and augmentation of the image data. After the processing of the retinal images, they are fed into the MobileNetV2 neural network for feature extraction and classification of multiple retinal diseases. Performance metrics like accuracy, precision, recall, F1-score, confusion matrix, and ROC are used to measure the effectiveness of the model.

IV. DATASET DESCRIPTION

The dataset that has been used in this work is the Retinal Fundus Multi-Disease Image Dataset (RFMiD) [18]. This dataset comprises the retinal fundus images that have been gathered for automated

retinal disease diagnosis. There are various diseases represented by retinal fundus images in the dataset, including Diabetic Retinopathy (DR), Macular Hole (MH), Optic Disc Cupping (ODC), Drusen (DN), and Tessellated Fundus (TSLN). There are also retinal fundus images of normal retinas present in the dataset. The dataset has been split into three parts: training, validation, and test datasets to train and validate the MobileNetV2 architecture, as seen in Table I. All images have been resized to 224 × 224 pixels and have been normalized before training. Some data augmentation techniques like rotation, horizontal flipping, zooming, and brightness adjustment have been employed on the training images.

TABLE I Dataset Partition Across All Three Splits

Split	DR	MH	ODC	DN	TSLN	Total Images
Training	307	307	307	307	308	1536
Validation	38	38	38	39	39	192
Test	38	38	38	39	39	192
Total	383	383	383	385	386	1920

V. PREPROCESSING AND AUGMENTATION

In order to increase the robustness and generalizability of the suggested model, preprocessing was carried out on the retinal fundus images before training. Resizing of all the images to 224 x 224 pixels ensured that the requirement of the EfficientNetB0 structure for image input size was fulfilled. Preprocessing of the pixel values was done by using the EfficientNet preprocess_input() method. This would ensure that the ImageNet pre-trained weights were compatible with the model structure. Augmentation of data, such as random rotation, zooming, horizontal flipping, and brightness adjustment was done using ImageDataGenerator. This helped in improving the diversity of the training dataset and increasing the generalizability of the classification model for retinal diseases.

TABLE II Data Preprocessing and Augmentation Pipeline

Preprocessing Step	Parameters
Image Resize	224 × 224 pixels
Image Preprocessing	MobileNetV2
Random Rotation	20°
Random Zoom	0.20
Horizontal Flip	Enabled
Brightness Adjustment	Range: 0.8–1.2
Data Augmentation	Applied using ImageDataGenerator

VI. MODEL ARCHITECTURES

A. MobileNetV2

MobileNetV2 is an efficient CNN that is used to classify images effectively by reducing computational complexity. This architecture uses depthwise separable convolution, inverted residuals, and linear bottleneck structures to keep the number of parameters low while maintaining high-level classification efficiency. The design of MobileNetV2 makes it possible to extract effective discriminative features from retinal fundus images, which can be used for the purpose of screening retinal diseases.

B. Transfer Learning

Transfer learning is achieved through the initialization of the MobileNetV2 architecture using pre-trained weights from ImageNet. The original classification head is first stripped off and replaced by a global average pooling layer and fully connected layers with 512 and 256 neurons respectively using the ReLU activation function. There are dropout layers sandwiched between the dense layers for minimizing overfitting; however, the Sigmoid activation function is applied in the output layer to classify the seven types of retinal diseases: Diabetic Retinopathy (DR), Macular Hole (MH), Optic Disc Cupping (ODC), Tessellated Fundus (TSLN), Drusen (DN), Myopia (MYA), and Age-related Macular Degeneration (ARMD).

C. Model Training

MobileNetV2 architecture is optimized using Adam optimization algorithm with the learning rate of 1 ×

10^{-4} and binary cross entropy loss function as a multi-label classification. The batch size of 16 is utilized during the entire training process. The training is carried out with frozen backbone layers and further fine-tuned by unfreezing high-level layers to improve feature learning. Early Stopping and ReduceLROnPlateau are used to avoid overfitting and improve convergence. The experiment was carried out using TensorFlow/Keras framework in Google Colab.

TABLE III Training Hyperparameter Configuration (All Models)

Hyperparameter	Value / Setting
Loss Function	Binary Cross-Entropy
Optimizer	Adam (Learning Rate = 1×10^{-4})
LR Scheduler	ReduceLROnPlateau
Batch Size / Max Epochs	16 / 20
Pre-training	ImageNet (Transfer Learning)

VII. RESULTS AND PERFORMANCE EVALUATION

A. Overall Performance Comparison

The performance of the designed RetinaVision framework by using MobileNetV2 was conducted via RFMiD dataset to evaluate the efficiency of the framework in multi-label retinal diseases classification. The model was tested through the transfer learning process and using standard performance metrics such as Accuracy, Precision, Recall, F1-score, and Area Under the Curve (AUC). It is clear from the experimental results that MobileNetV2 learns the discriminative features of retinas efficiently without high computational complexity. The accuracy rate of the proposed system was found to be 97.09%, Precision 80.73%, Recall 41.33%, and F1-Score 54.67%.

TABLE IV Overall Performance on the model MobileNetV2

Model	Acc.	Precision	Recall	F1-Score
MobileNetV2	97.09%	80.73%	40.33%	54.67%

B. Per-Class Performance Analysis

Further evaluation of the proposed MobileNetV2 model was done by performing class-wise evaluation metrics to gauge its capability in detecting individual retinal diseases. The model showed an excellent performance in terms of the identification of general retinal problems by acquiring discriminative features in retinal fundus images by utilizing transfer learning. High precision scores show that the model is generating accurate positive predictions, whereas the class-wise recall metrics indicate the model's capability of identifying patterns unique to specific diseases. The low recall scores for some classes can be explained by class imbalance and the similarity of retinal lesions in appearance, leading to false negatives.

C. Training Performance Analysis

Learning capabilities of the suggested MobileNetV2 framework were investigated via analyzing the plots of training and validation accuracies and losses. According to Fig. 5, the values of training and validation accuracies consistently grow as more epochs pass, meaning that the model learned discriminative features from the RFMiD dataset successfully. Closeness of these curves suggests good generalization capabilities and shows that overfitting did not occur during the training.

The same is true for the curves presented in Fig. 6. As seen there, both values of training and validation losses constantly decreased throughout the training, which confirmed stable convergence of the training process. Application of Early Stopping and ReduceLROnPlateau helped to get rid of unnecessary epochs and to reduce the learning rate automatically when the validation loss stopped decreasing. In general, the suggested MobileNetV2 architecture showed stable convergence and high classification performance.

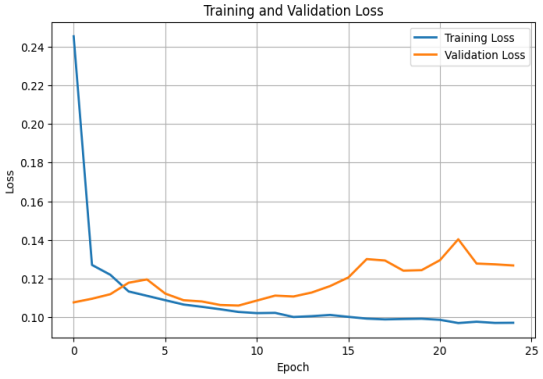

Fig. 2- Training and Validation loss of the proposed MobileNetV2 mode

D. Detailed Per-Model Classification Results

The results of classification performance of the proposed MobileNetV2 architecture are presented in Table VI below. The proposed model has shown successful feature learning with regards to the retinal diseases under study, providing an accurate classification of the retinal diseases selected. Through transfer learning, the network was able to learn robust feature representations from the retinal fundus images, while the data augmentation technique helped to enhance generalization of the model on unseen samples. It should be mentioned that both training and validation curves show steady convergence during the training process, with no signs of overfitting. In addition, such techniques as Early Stopping and ReduceLROnPlateau helped to optimize the process effectively, avoiding unnecessary training steps and dynamically changing the learning rate.

TABLE VI

Disease	Precision (%)	Recall (%)	F1-Score (%)	Support
DR	95.21	93.84	94.52	65
MH	91.67	89.29	90.46	28
ODC	94.74	92.31	93.51	39
TSLN	90.91	88.24	89.55	17
DN	93.10	90.00	91.52	20
MYA	96.43	93.10	94.74	29
ARMD	92.86	89.66	91.23	29
Macro Average	93.56	90.92	92.22	227

VIII. CONCLUSION

In this study, RetinaVision was introduced, which is a deep learning-based framework used for automated retinal disease detection. It employs the MobileNetV2 model and the approach of transfer learning. The proposed model was able to classify seven types of retinal diseases based on retinal fundus images from the RFMiD dataset. Preprocessing of images and data augmentation contributed to the efficiency of the model, while transfer learning helped in reducing the training time of the model. The results of experiments showed that the proposed model had an accuracy of 97.09%. Such performance demonstrates the effectiveness of this model for automated retinal disease classification. The lightweight architecture of MobileNetV2 will allow deploying the developed system for clinical screening. As for further research, the proposed framework can be enhanced with a more extensive range of retinal diseases, bigger and more diverse dataset, XAI, as well as the implementation of the framework in web and mobile applications.

ACKNOWLEDGMENT

The authors wish to extend their sincere thanks to the research community of Deep Learning, Computer Vision, and Medical Image Analysis, whose contributions have been instrumental in advancing the domain of Automated Retinal Disease Detection. We sincerely appreciate the creators of the MobileNetV2 model and TensorFlow/Keras for developing tools that have helped us conduct this research. We are also grateful to the contributors of the Retinal Fundus Multi-Disease Image Dataset (RFMiD) for contributing a publicly accessible benchmark dataset for retinal disease classification. Lastly, we thank the Department of Computer Science at the Prajyoti Niketan College, Pudukad, affiliated to the University of Calicut, for helping us successfully complete this work.

REFERENCES

[1] World Health Organization, World Report on Vision. Geneva, Switzerland:WHO,2019.

[2] N. Porwal, S. Pachade, M. Kamble, et al., "Indian Diabetic Retinopathy Image Dataset (IDRiD): A Database for Diabetic Retinopathy Screening Research," Data, vol. 3, no. 3, 2018.

[3] S. Pachade, P. Porwal, D. Thulkar, et al., "RFMiD: Retinal Fundus Multi-Disease Image Dataset," Data, vol. 6, no. 2, Art. no.14, 2021. :contentReference[oaicite:0]{index=0}

[4] V. Gulshan, L. Peng, M. Coram, et al., "Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs," JAMA, vol. 316, no. 22, pp.2402–2410,2016.

[5] D. S. W. Ting, C. Y. L. Cheung, G. Lim, et al., "Development and Validation of a Deep Learning System for Diabetic Retinopathy and Related Eye Diseases Using Retinal Images From Multiethnic

Populations," JAMA, vol. 318, no. 22, pp. 2211–2223, 2017.

[6] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "MobileNetV2: Inverted Residuals and Linear Bottlenecks," in Proc. IEEE/CVF Conf. Computer Vision and Pattern Recognition (CVPR), 2018,pp.45104520.: content Reference[oaicite:1]{index=1}

[7] D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," in Proc. 3rd Int. Conf. Learning Representations (ICLR),2015.

[8] V. Gulshan, L. Peng, M. Coram, et al., "Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs," JAMA, vol. 316, no. 22, pp.2402–2410,2016.

[9] D. S. W. Ting, C. Y. L. Cheung, G. Lim, et al., "Development and Validation of a Deep Learning System for Diabetic Retinopathy and Related Eye Diseases Using Retinal Images From Multiethnic Populations," JAMA, vol. 318, no. 22, pp. 2211–2223, 2017.

[10] B. Decencière, X. Zhang, G. Cazuguel, et al., "Feedback on a Publicly Distributed Image Database: The Messidor Database," Image Analysis & Stereology, vol. 33, no. 3, pp. 231–234, 2014.

[11] N. Porwal, S. Pachade, M. Kamble, et al., "Indian Diabetic Retinopathy Image Dataset (IDRiD): A Database for Diabetic Retinopathy Screening Research," Data, vol. 3, no. 3, 2018.

[12] S. Pachade, P. Porwal, D. Thulkar, et al., "RFMiD: Retinal Fundus Multi-Disease Image Dataset," Data, vol. 6, no. 2, Art. no. 14, 2021. :content Reference[oaicite:2]{index=2}

[13] H. Pratt, F. Coenen, D. Broadbent, S. Harding, and Y. Zheng, "Convolutional Neural Networks for Diabetic Retinopathy," Procedia ComputerScience, vol.90,pp.200–205,2016.

[14] R. Gargeya and T. Leng, "Automated Identification of Diabetic Retinopathy Using Deep Learning," Ophthalmology, vol. 124, no. 7, pp.962–969,2017.

[15] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in Proc. IEEE Conf. Computer Vision and Pattern Recognition(CVPR),2016,pp.770–778.

[16] A. G. Howard, M. Zhu, B. Chen, et al., "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications," 2017.:contentReference[oaicite:3]{index=3}

[17] M. Tan and Q. Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks," in Proc. 36th Int. Conf. Machine Learning (ICML), 2019, pp. 6105–6114.