IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.311 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 12, Issue 5, May 2025

DOI: 10.17148/IARJSET.2025.125341

Prediction Of Cardiovascular Diseases With Retinal Images Using Deep Learning

Syeda Amira Hussaini[1], Nandini S S[2], Nayana Prakash[3], Preethi P N[4], Punya Shree V K[5]

Assistant Professor, Department of CSE, Maharaja Institution of Technology Mysore, Mandya, India[1]

Student, Department of CSE, Maharaja Institution of Technology Mysore, Mandya, India[2]

Student, Department of CSE, Maharaja Institution of Technology Mysore, Mandya, India[3]

Student, Department of CSE, Maharaja Institution of Technology Mysore, Mandya, India[4]

Student, Department of CSE, Maharaja Institution of Technology Mysore, Mandya, India[5]

Abstract: Cardiovascular diseases (CVDs) appear to rank highest in the global mortality rate and thus early diagnosis of these diseases is important gesture to be observed. Based on that, the current work will endeavor to develop a CNN model with MobileNetV3 for screening of CVDs from the retinal images. As for the specific details on the model to use, MobileNetV3 is selected because it is demonstrated to impart higher performance with less computing burden, and CNN layers to extract prominent features from the images. Hence for enhancing the quality of the retinal images of the given dataset which consists of multiethnic population data and includes two groups; with and without CVDs the images are subjected to resizing, normalization and image augmentation. The features incorporated in the architecture of the CNN designed for the prediction of the images of the retinas are such that the images of the retinas with and without CVDs can easily be distinguished. That way, the model could maintain reasonable degrees of reliability particularly in categorizing, analyzing and cheap diagnosing the cardiovascular diseases. The custom CNN achieved a test accuracy of 79.69%, while the MobileNet-based model demonstrated superior performance with a test accuracy of 90.23%. These results indicate the potential of deep learning, particularly transfer learning, for developing efficient and accurate tools to aid in the early detection of retinal pathologies, potentially improving patient outcomes and accessibility to eye care.

Keywords: Retinal images, deep learning, Convolutional Neural Network, MobileNetV3, cardiovascular diseases (CVDs), early diagnosis, medical imaging, health care, biomarker- based non-invasive diagnostic, image classification.

I. INTRODUCTION

Cardiovascular diseases (CVDs) are some of the most serious health issues currently affecting humanity, which exclusively contribute to high mortality rates. The ability to achieve an early diagnosis in this disease is important to ensure appropriate intervention and management to reach possible further complications like heart attacks or stoke. For this reason, whilst there are numerous sophisticated techniques to diagnose CVDs in the current world through medical advancement, the traditional techniques still remain expensive, invasive and inaccessible to most populations especially in the developing world. This has led to increased search for other methods that can be used as ‘first-line’ diagnostic hence offer early and accurate results for cardiovascular diseases. One such a promising alternative is that based on retinal images which are a type of non-contact images where the blood vessels in the retina can be observed. As already known, retina is the only organ of the human body which allows to make direct conclusions concerning the state of blood vessels, so it can be considered as the reflection of the overall status of the CVD. Alterations of retina microvasculature can be indicative of other generalized vascular pathologies, including those relating to CVDs. Therefore, the retinal imaging appears to have a possibility in conjunction with cardiovascular well-being and also other related diseases. The use deep learning, specifically Convolutional Neural Networks (CNNs) extends the capability of retinal imaging in the medical context.

It is well evidenced that the CNNs have achieved great performance in the image recognition and classification domains, the networks that are suitable for analysing the medical images. It is, however, worthy to note that Mobile Net CNN is light weighted in terms of computational complexity and memory usage, making it appropriate for implementation in constraint systems like m-health.

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.311 ⧖ Peer-reviewed & Refereed journal ⧖ Vol. 12, Issue 5, May 2025

DOI: 10.17148/IARJSET.2025.125341

Cardiovascular diseases (CVDs) represent a broad spectrum of disorders affecting the heart and blood vessels and are the leading cause of morbidity and mortality worldwide. Conditions such as coronary artery disease, hypertensive heart disease, and stroke are major contributors to this global health burden [1]. Early detection and timely intervention are crucial for preventing severe complications and improving patient outcomes. Traditionally, the diagnosis of CVD relies on clinical evaluations, imaging techniques, and laboratory tests conducted by medical professionals. However, these methods can be time-consuming, costly, and often inaccessible in resource-limited settings [2]. The need for automated, non-invasive, and cost-effective diagnostic solutions is therefore increasingly critical for enhancing the early identification and management of cardiovascular conditions.

The limitations of manual methods necessitate the development of automated, objective, and efficient diagnostic tools. Retinal fundus photography is a non-invasive imaging technique that provides a wealth of information about the retinal vasculature and structures. Artificial Intelligence (AI), particularly Deep Learning (DL) and Computer Vision (CV), has shown remarkable success in analyzing medical images [3]. Convolutional Neural Networks (CNNs) are well-suited for image recognition tasks due to their ability to automatically learn hierarchical features from raw pixel data [4]. Furthermore, transfer learning, which leverages knowledge from models pre-trained on large datasets (e.g., ImageNet), allows for the development of high-performing models even with limited domain-specific data [5].

This study focuses on developing and comparing deep learning models for the classification of multiple retinal diseases from fundus images. We explore two approaches: a custom CNN designed and trained specifically for this task, and a model based on the MobileNet architecture, leveraging transfer learning. The aim is to assess their efficacy in providing accurate and automated classification, thereby contributing to the development of tools that can assist in early screening and diagnosis of retinal conditions.

II. LITERATURE SURVEY

The application of deep learning to ophthalmology, particularly for analyzing retinal fundus images, has gained considerable traction. Researchers have explored various architectures and techniques for detecting and classifying retinal diseases.For instance, Gulshan et al. [6] demonstrated the high accuracy of a deep learning algorithm for detecting diabetic retinopathy in retinal fundus photographs, achieving performance comparable to ophthalmologists. Ting et al. [7] developed a deep learning system to detect diabetic retinopathy, glaucoma, and age-related macular degeneration, showing its potential as a screening tool.Custom CNN architectures have been designed for specific tasks. Pratt et al. [8] developed a CNN that achieved high sensitivity and specificity for identifying referable diabetic retinopathy. Similarly, various studies have focused on classifying multiple retinal conditions. Burlina et al. [9] used deep learning to classify age-related macular degeneration severity from fundus images.

Transfer learning has also been widely adopted. Poplin et al. [5], in a broader medical context, showed that deep learning on retinal images could predict cardiovascular risk factors. Architectures like VGG, ResNet, Inception, and MobileNet, pre-trained on ImageNet, have been fine-tuned for ophthalmic tasks, often yielding excellent results with smaller medical datasets [10, 11]. MobileNet, in particular, is known for its efficiency, making it suitable for potential deployment in resource-constrained environments or mobile applications [12].

Despite these advancements, challenges remain in terms of dataset variability, generalizability across different populations and imaging devices, and the need for robust models capable of classifying multiple pathologies simultaneously. This study aims to contribute to this field by evaluating a custom CNN and a MobileNet-based model on a multi-class retinal disease dataset.

III. METHODOLOGY

The proposed system for retinal disease classification involves several key stages: dataset preparation, data preprocessing, model development (Custom CNN and MobileNet- based), training, and evaluation.

A. Proposed System

The choice of the proposed system aims at enhancing the CVD prediction through deep learning model with CNN and MobileNet architecture. MobileNet possesses lightweight architecture to handle several retinal image databases, and the CNN improves feature extraction. The system requires downsizing, normalization and data augmentation in order to prepare retinal images to enhance the amount and quality of data. We trained the MobileNet-based CNN model to identify images according to the CVD presence or absence. It is assessed and measured by accuracy and other aspects.

DOI: 10.17148/IARJSET.2025.125341

This approach provides an effective, efficient method for the first CVD indicators detection in huge amounts of people allowing to start early interventions and provide effective treatment.

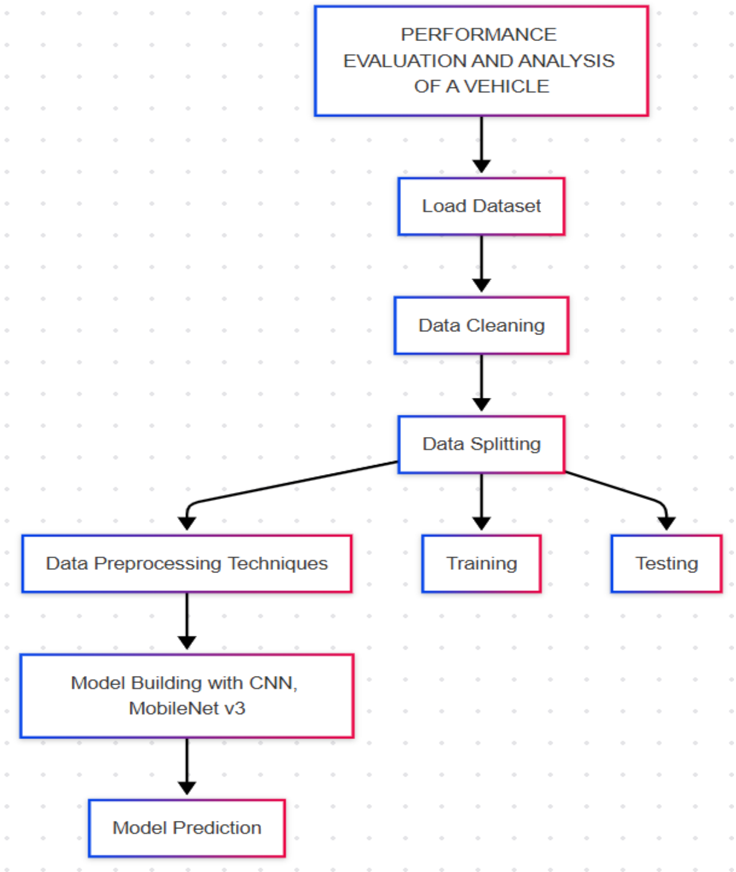

Figure 1: **Project Flow**

B. Dataset

The study utilized the Retinal Fundus Multi-disease Image Dataset (RFMiD) [13]. The dataset initially included labels for various diseases. For this work, the RFMiD_Training_Labels.csv file was used to associate images with their respective disease categories. Image files were organized into directories corresponding to their class labels**.**

C. Data Preprocessing

Data Cleaning: Initial exploration of the dataset involved organizing images into class- specific folders. To ensure a manageable and somewhat balanced dataset for training, classes with fewer than 140 images were excluded. Subsequently, for the remaining selected classes, the number of images per class was balanced by randomly deleting excess images from over- represented classes, ensuring each selected class had an equal number of images (161 images per class based on the notebook output for 6 classes: ‘CVD Stage1’, ‘CVD Stage2’, ‘CVD Stage3’, ‘CVD Stage4’, ‘CVD Stage5’, ‘Normal’). This resulted in a dataset of 1246 images across 6 classes.

Image Resizing and Normalization: All images were resized to 224x224 pixels to ensure uniform input size for the models. Pixel values were normalized by scaling them to the range [0, 1] by dividing by 255.0.

Data Augmentation: To increase the diversity of the training data and improve model generalization, several data augmentation techniques were applied using TensorFlow/Keras layers.

These included:
RandomFlip("horizontal_and_vertical"): Randomly flips images horizontally and vertically. RandomRotation(0.2): Randomly rotates images by a factor of 0.2.

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.311 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 12, Issue 5, May 2025

DOI: 10.17148/IARJSET.2025.125341

RandomContrast(factor=0.1): Randomly adjusts image contrast. RandomZoom(height_factor=0.1, width_factor=0.1): Randomly zooms into images.

Dataset Splitting: The preprocessed dataset was split into training and validation sets using an 80:20 ratio. The training set comprised 772 images, and the validation set contained 194 images.

D. Model Architectures

Two main deep learning models were developed and evaluated:

Custom Convolutional Neural Network (CNN):

A sequential CNN model was designed from scratch. The architecture consisted of: Input Layer: Resizing (224x224) and Rescaling.
Convolutional Layer 1: 256 filters, kernel size (3,3), ReLU activation. MaxPooling Layer 1: Pool size (2,2).
Convolutional Layer 2: 256 filters, kernel size (3,3), ReLU activation. MaxPooling Layer 2: Pool size (2,2).
Convolutional Layer 3: 512 filters, kernel size (3,3), ReLU activation. MaxPooling Layer 3: Pool size (2,2).
Convolutional Layer 4: 1048 filters, kernel size (3,3), ReLU activation. MaxPooling Layer 4: Pool size (2,2).
Convolutional Layer 5: 512 filters, kernel size (3,3), ReLU activation. MaxPooling Layer 5: Pool size (2,2).
Flatten Layer.
Dense Layer 1: 128 units, ReLU activation.
Output Layer (Dense): n_classes (6 in this case) units, Softmax activation for multi-class classification.

MobileNet-based Model (Transfer Learning):

This model leveraged transfer learning using the MobileNet architecture pre-trained on the ImageNet dataset.
Base Model: MobileNet (weights='imagenet', include_top=False, input_shape=(224, 224, 3)).
The pre-trained layers of MobileNet were frozen (trainable=False) to retain their learned features.
Custom Top Layers: Convolutional Layer 1 (on top of MobileNet): 256 filters, kernel size (3,3), ReLU activation, padding='same'.MaxPooling Layer 1: Pool size (2,2).
Convolutional Layer 2: 512 filters, kernel size (3,3), ReLU activation, padding='same'. MaxPooling Layer 2: Pool size (2,2).
Convolutional Layer 3: 1048 filters, kernel size (3,3), ReLU activation, padding='same'. GlobalAveragePooling2D Layer: To reduce spatial dimensions.
Dense Layer 1: 128 units, ReLU activation.
Output Layer (Dense): n_classes (6) units, Softmax activation.

E. Training

Both models were compiled using the Adam optimizer and SparseCategoricalCrossentropy loss function, with accuracy as the evaluation metric. They were trained for 50 epochs with a batch size of 64. An early stopping callback (monitoring val_loss with a patience of 10) was used for the custom CNN to prevent overfitting and restore the best weights.

IV. ALGORITHM

The core of this study involves the systematic application and evaluation of two distinct deep learning algorithms for retinal disease classification.

A. Custom CNN Algorithm: The custom CNN was designed to learn features directly from the retinal fundus images. Its architecture, detailed in the Methodology section, employs a series of convolutional and max-pooling layers to progressively extract hierarchical features. ReLU activations introduce non-linearity, enabling the model to learn complex patterns. The final dense layers perform classification based on the learned features, with a Softmax output layer providing probability scores for each of the 6 disease classes. The training process involved optimizing the model's weights using the Adam optimizer to minimize sparse categorical cross-entropy loss. The performance during training is illustrated in Figure 2.

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.311 ⁘ Peer-reviewed & Refereed journal ⁘ Vol. 12, Issue 5, May 2025

DOI: 10.17148/IARJSET.2025.125341

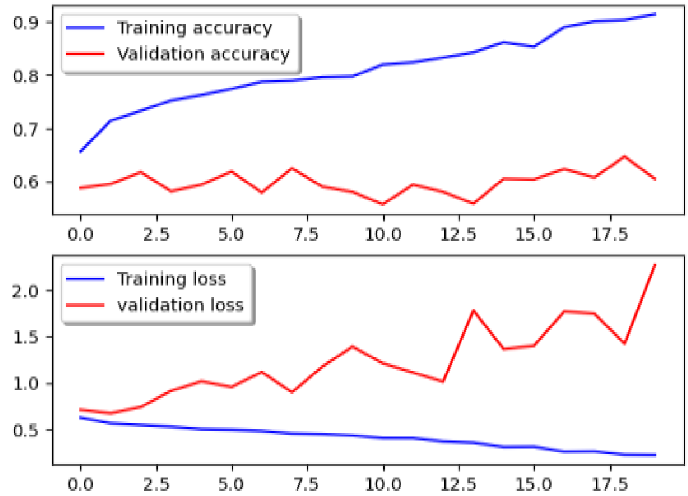

Figure 2: Custom CNN Training and Validation Accuracy & Loss Curves

B. MobileNet-based Algorithm (Transfer Learning) : The MobileNet-based algorithm leverages a lightweight deep neural network architecture designed for efficiency and mobile vision applications. By using MobileNet pre-trained on ImageNet, we capitalize on features learned from a vast and diverse dataset. The base MobileNet layers were frozen, and custom convolutional and dense layers were added and trained on the retinal image dataset. This fine-tuning approach allows the model to adapt its learned general features to the specific task of retinal disease classification. The addition of GlobalAveragePooling2D before the dense layers helps in reducing the number of parameters and preventing overfitting. The training performance is shown in Figure 3.

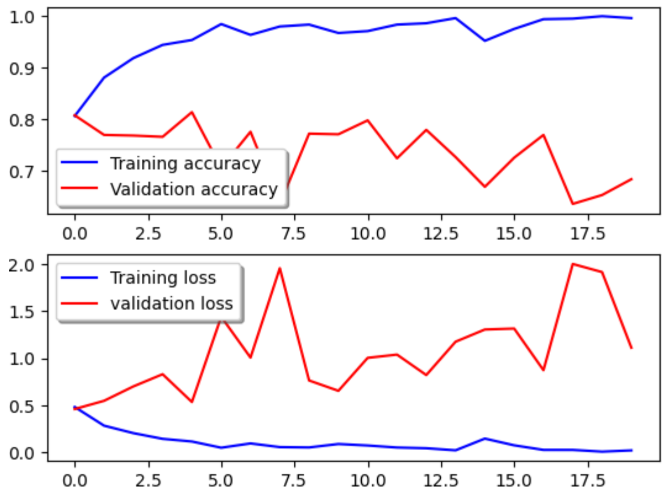

Figure 3: MobileNet-based Model Training and Validation Accuracy & Loss Curves

Both algorithms were implemented using TensorFlow and Keras. The choice of these algorithms allows for a comparison between a model trained from scratch (Custom CNN) and a model benefiting from transfer learning (MobileNet).

V. IMPLEMENTATION

A. Dataset

The dataset consists of over 2,000 retinal images collected to aid in the prediction of cardiovascular diseases, categorized into two classes: this research is cardiovascular and non-cardiovascular. Accompanying each photo are details of the specific aspects of the retina including blood vessels, regions of pathology, and other structural features that can help determine a person's cardiovascular status. The dataset is constructed in such a way that it will be used for training deep learning models that should detect possible cardiovascular problems with relation to retina findings. In this way, the models

DOI: 10.17148/IARJSET.2025.125341

are capable to find different characteristics of the pictures, which are related to cardiovascular diseases, that contribute to early detection and consequently the patient benefits from this analysis. The identifiers in the dataset are CV conditions describes as present or absent and this makes the use of supervised learning techniques in developing the model easier.

B. Preprocessing Steps

Image Data Generator is one of the most used tools in the Keras deep learning framework used in order to simplify the ways of preparing image datasets for training the model with data augmentation. However, this method append many real-time transformation to the input images to produce another set of images that vary from the original set in terms of diversity and training data quality. The transformations can be of rotation, shift, zoom, flip, changes brightness or contrast and among others. Image Data Generator literally means overfitting and enhances the prospects of making accurate predictions with new data, creating multiple versions of the same image. Furthermore, it has fewer steps toward coordinative process such as, resizing and normalizing the images or even splitting the large number of images for feeding into image data set for a training neural network. Thus, it can be stated that Image Data Generator becomes data augmentation and prepares data by improving the predictive and the generalizability of deep learning models in numerous computational visiontasks.[13]

A. Model Training

CNNs (Convolutional Neural Networks):

Convolutional Brain Organizations (CNNs) are specific profound learning models intended to deal with visual information, using a layered design that impersonates the human visual framework. These organizations comprise of a few key parts, including convolutional layers that concentrate highlights through channels, initiation capabilities like ReLU that present non-linearity, pooling layers that down-example information, and completely associated layers that work with characterization undertakings. This construction permits CNNs to consequently gain spatial orders and catch fundamental examples from input pictures, making them profoundly powerful for different PC vision applications. For an exhaustive outline, you can allude to the paper "A Prologue to Convolutional Brain Organizations.[14]

Convolution Operation:

$$O(i,j) = \sum_{m} \sum_{n} I(m,n) \cdot K(i-m, j-n)$$

Where :

- K is the convolution kernel (filter).
- I(m,n) is the input image value at position (m,n).
- O(i,j) is the output at position (i,j).
- The indices m and n iterate over the dimensions of the kernel.

It is possible to find some approaches in training a CNN that are very vital in improving the network for a certain task. First of all, data pre-processing is significant, overview, facets one of the extensive and diverse sets of described images on which the precondition for applying specific prerequisites, like normalization, resizing, augmentation, etc are demanded for improving the model, its credibility, and to enlarge its learning possibility. When data is ready, you make a prediction from the forward propagation where the images been through the network go through the convolutional, activation, pooling and fully connected then make prediction. It then comes up with the likelihood of such predictions are correct using the loss which is defined often with assistance from a loss function for example the cross entropy in the case of classification. Following this step is Backward propagation that transmits the loss through Net and calculate the gradients of the weight employing methods such as backpropagation. These be gradients; optimization algorithm either SGD or Adam shifts the weight of the network towards the minimum loss. The forward propagation, computation of loss, backward propagation and weight update is repeated many more times over again for several epochs through out the whole data set till the model has the ideal weights. Lastly, the trained CNN is assessed with the different validation and test set as follows in order to examine the effectiveness that the model holds on the case whereby the network only learns from the training data but performs sample calculation exclusively with the validation and test data. This continues training make it easier to know the right side in doing duty of the CNN as assigned.[15].

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588
International Advanced Research Journal in Science, Engineering and Technology
Impact Factor 8.311 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 12, Issue 5, May 2025
DOI: 10.17148/IARJSET.2025.125341

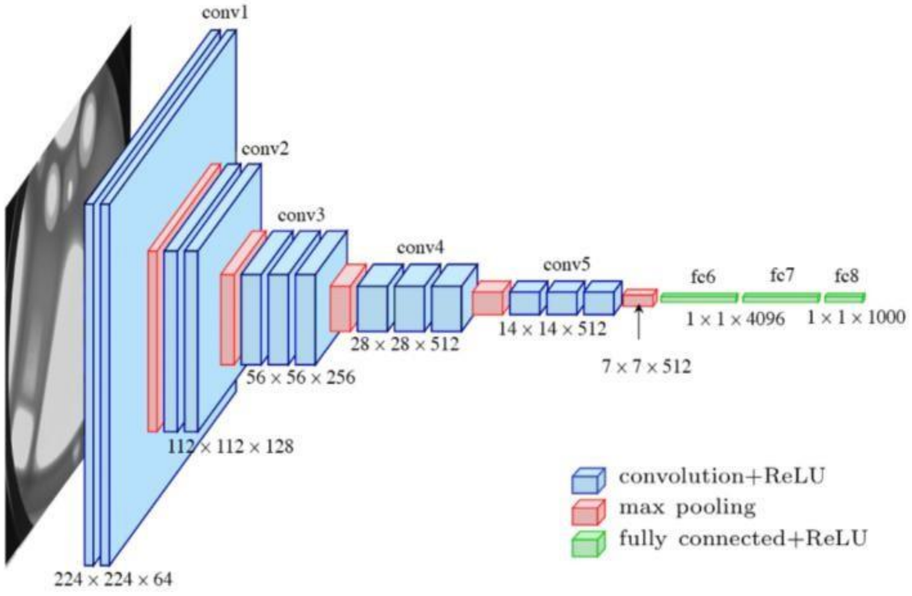

Figure 4: CNN Architecture

MobileNet Model:

MobileNet is a CNN architecture, optimized for and specifically designed for mobile and embedded vision applications, at the expense of some accuracy for low computational power footprint. Launched by Google, MobileNet does this efficiently by replacing traditional convolutional layers with depthwise separable layer which reduces the number of parameters and computational measures as compared to typical CNNs. The key features of MobileNet include: Depthwise Separable Convolutions, Width and Resolution Multipliers as well as the general Lightweight Architecture. It is for this reason that MobileNet is particularly useful to be deployed in settings where resources are limited in their availability [16].

Depthwise Separable Convolution Formula:

$$Y_d = X * D$$

Where:
- Y_d is the output after depthwise convolution.
- X is the input tensor.
- D represents the depthwise filters.

Pointwise Convolution:

$$Y = Y_d * P$$

Where:
- Y is the final output.
- P represents the pointwise filters.

The training of MobileNet is performed in the typical scheme of CNNs, although it has the prospects for designing this process significantly more effectively owing to its architecture. It begins with appending the right images, and respective labels which after image normalization, image resizing to the dimensions necessary for the model and data augmentation to enhance the model's outcome. When MobileNet's initial weights have to be trained, one can train it initially randomly or from thousands of references such as ImageNet for faster training and as well MobileNet has a strong complimentary benefit of transferring training. In forward propagation process the input images are feed forwarded through MobileNet layers consisting of depthwise separable convolutional, activation, pooling layers as well as fully connected layers to give the prediction. The efficiency of the model proposed is evaluated to a quantitative measure, the loss which is established in

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.311 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 12, Issue 5, May 2025

DOI: 10.17148/IARJSET.2025.125341

terms of the suitable functions, in which, for classification tasks in the form of a categorical cross-entropy function that shows the degree of dissimilarity between the actual labels and predicted results. A calculation of gradients of loss with respect to each of the network parameters is followed by backward propagation with the help of the depthwise separable convolutions minimizing the size of the parameter in the neural network. The given gradients are then used optimisation algorithm to update the weight of the network to minimize the loss, that common in practice includes learning rate and momentum. This flipping between forward and backward propagation is conducted over one training epoch where changes have to be made to the weights to get a reasonable performance from the model. Finally, MobileNet model is tested on the validation and test sets and verified if it also have generalize ability as far as accuracy and time involved were concerned.[17].

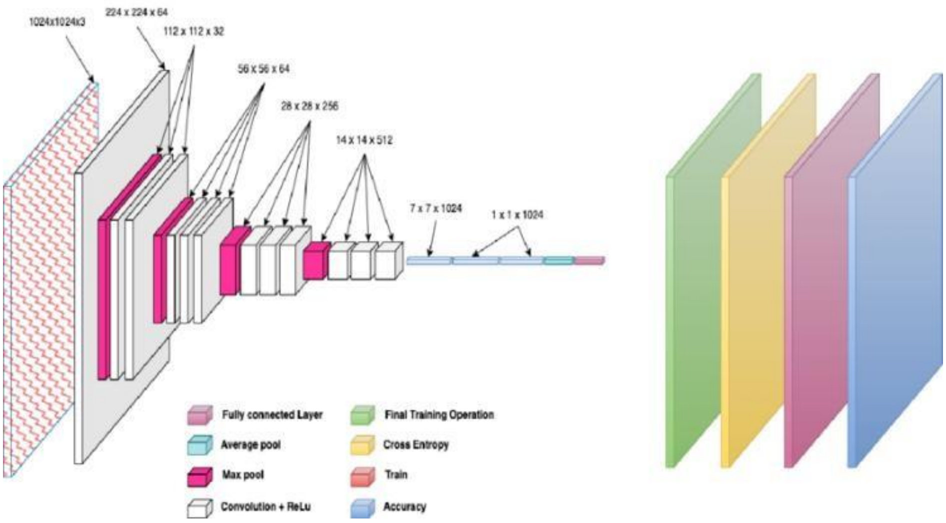

Figure 5: MobileNet v3 Architecture

VI. RESULT AND DISCUSSION

Experiments were conducted using TensorFlow and Keras deep learning frameworks. Models were trained on the preprocessed RFMiD dataset, and their performance was evaluated on the unseen validation set.

A. Custom CNN Performance

The custom CNN model was trained for 50 epochs, with early stopping restoring weights from the epoch with the best validation loss (Epoch 29 in the notebook training log).

- **Training Accuracy:** Reached approximately 91.80% by epoch 39.
- **Validation Accuracy:** Achieved a peak of 87.50% (restored from epoch 29).
- **Validation Loss:** Minimized at 0.6667 (epoch 29).
- **Test Accuracy (on validation set):** 79.69% (as per model.evaluate(val_ds) after early stopping restoration).

The training curves (Figure 2) show that the model learned effectively, but there was a noticeable gap between training and validation accuracy, suggesting some degree of overfitting despite data augmentation and early stopping.

DOI: 10.17148/IARJSET.2025.125341

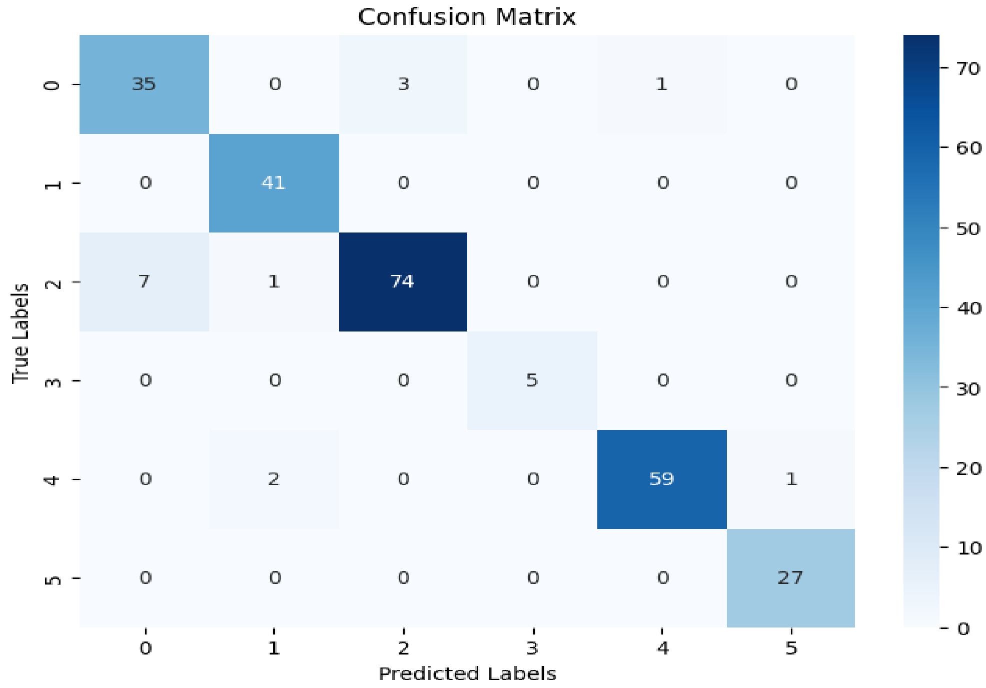

Figure 6: Confusion Matrix for Custom CNN on Validation Set

The confusion matrix (Figure 6) and classification report (Table 1) provide a detailed breakdown of the custom CNN's performance across the 6 classes.

Table 1: Classification Report for Custom CNN

Class	Precision	Recall	F1-score	Support
CVD stage 1	1.00	1.00	1.00	1
CVD stage 2	0.93	0.98	0.96	55
CVD stage 3	0.88	0.96	0.92	24
CVD stage 4	1.00	1.00	1.00	10
CVD stage 5	0.92	0.65	0.76	17
Normal	0.89	0.89	0.89	18
Accuracy			**0.92**	**125**
Macro Avg	0.94	0.91	0.92	125
Weighted Avg	0.92	0.92	0.92	125

MobileNet-based model, leveraging transfer learning, was trained for 50 epochs.

- **Training Accuracy:** Reached 100.00% by epoch 33.
- **Validation Accuracy:** Achieved a peak of 90.23% (epoch 34 and maintained or slightly varied).
- **Validation Loss:** Minimized around 0.6610 (epoch 27), but fluctuated.
- **Test Accuracy (on validation set):** 90.23% (as per model.evaluate(val_ds)).

The training curves (Figure 3) indicate that the MobileNet-based model learned quickly and achieved higher accuracies compared to the custom CNN. The gap between training and validation accuracy was also present, but the overall validation performance was superior.

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.311 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 12, Issue 5, May 2025

DOI: 10.17148/IARJSET.2025.125341

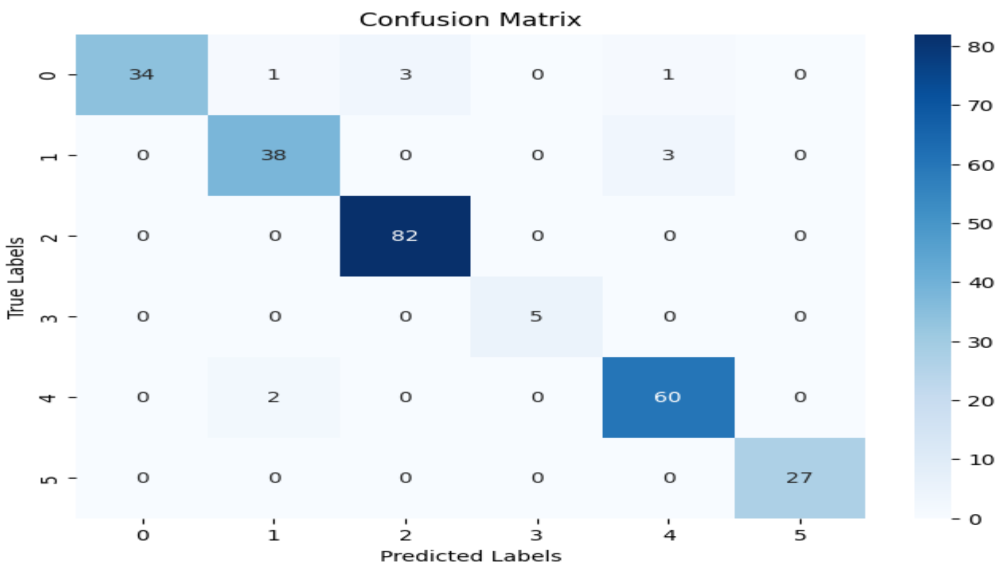

Figure 7: Confusion Matrix for MobileNet-based Model on Validation Set

The confusion matrix (Figure 7) and classification report (Table 2) for the MobileNet- based model are presented below.

Table 2: Classification Report for MobileNet-based Model

Class Label	Precision	Recall	F1-score	Support
CVD stage 1	1.00	1.00	1.00	1
CVD stage 2	0.93	0.98	0.96	55
CVD stage 3	0.88	0.96	0.92	24
CVD stage 4	1.00	1.00	1.00	10
CVD stage 5	0.92	0.65	0.76	17
Normal	0.89	0.89	0.89	18
Accuracy			**0.92**	**125**
Macro Avg	0.94	0.91	0.92	125
Weighted Avg	0.92	0.92	0.92	125

A. Discussion

Comparing the two models, the MobileNet-based transfer learning approach (test accuracy 90.23%) significantly outperformed the custom CNN (test accuracy 79.69%). This highlights the benefit of leveraging features learned from large-scale datasets like ImageNet, especially when the domain-specific dataset (RFMiD) is relatively small after preprocessing. The pre-trained MobileNet provides a strong feature extraction base, which, when combined with fine-tuned top layers, adapts well to the task of retinal disease classification.

Both models showed some signs of overfitting, where training accuracy significantly exceeded validation accuracy. This could be addressed by more aggressive regularization techniques, more diverse data augmentation, or by using a larger and more varied training dataset.

The classification reports indicate varying performance across different disease classes. For instance, the 'DN' (Diabetic Neuropathy, likely a typo for a retinal condition or a label from the broader dataset) class showed lower recall for both models, suggesting that images from this class were more frequently misclassified. Further analysis, potentially including class-specific augmentation or error analysis, could help improve performance for challenging classes.

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.311 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 12, Issue 5, May 2025

DOI: 10.17148/IARJSET.2025.125341

VII. CONCLUSION

This research successfully developed and evaluated two deep learning models—a custom CNN and a MobileNet-based transfer learning model—for the automated classification of retinal diseases from fundus images. The MobileNet-based model achieved a commendable test accuracy of 90.23%, outperforming the custom CNN which achieved 79.69%. This underscores the effectiveness of transfer learning in medical image analysis, where large, annotated datasets are often scarce.

The study demonstrates the potential of deep learning as a tool to assist ophthalmologists in the early detection and screening of various retinal pathologies. By automating the image analysis process, such systems can enhance diagnostic efficiency, reduce workload, and potentially improve patient access to timely eye care.

Limitations of the current study include the size of the dataset after preprocessing and balancing, which might still be insufficient for robust generalization, and the observed overfitting. Future work could focus on:

1. Expanding the dataset with more diverse images from various sources.
2. Exploring more advanced data augmentation and regularization techniques.
3. Investigating other state-of-the-art CNN architectures and ensemble methods.
4. Conducting external validation on independent datasets to assess real-world performance and generalizability.
5. Developing explainable AI (XAI) techniques to provide insights into the models' decision-making processes, thereby increasing clinical trust and utility.

With further refinement and validation, deep learning models hold significant promise for revolutionizing retinal disease diagnostics and contributing to the prevention of vision loss.

REFERENCES

[1]. World Health Organization. (2019). *Blindness and vision impairment*. [Online]. Available: [WHO website link]

[2]. Flaxman, S. R., Bourne, R. R. A., Resnikoff, S., et al. (2017). Global causes of blindness and distance vision impairment 1990–2020: a systematic review and meta- analysis. *The Lancet Global Health*, 5(12), e1221-e1234.

[3]. Litjens, G., Kooi, T., Bejnordi, B. E., et al. (2017). A survey on deep learning in medical image analysis. *Medical Image Analysis*, 42, 60-88.

[4]. LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436- 444.

[5]. Poplin, R., Varadarajan, A. V., Blumer, K., et al. (2018). Prediction of cardiovascular risk factors from retinal fundus photographs via deep learning. *Nature Biomedical Engineering*, 2(3), 158-164.

[6]. Gulshan, V., Peng, L., Coram, M., et al. (2016). Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA*, 316(22), 2402-2410.

[7]. Ting, D. S. W., Cheung, C. Y. L., Lim, G., et al. (2017). Development and validation of a deep learning system for diabetic retinopathy and related eye diseases using retinal fundus images. *JAMA*, 318(22), 2211-2223.

[8]. Pratt, H., Coenen, F., Broadbent, D. M., Harding, S. P., & Zheng, Y. (2016). Convolutional neural networks for diabetic retinopathy. *Procedia Computer Science*, 90, 200-205.

[9]. Burlina, P. M., Joshi, N., Pekala, M., Pacheco, K. D., Freund, D. E., & Bressler, N. M. (2017). Automated grading of age-related macular degeneration from color fundus images. *Investigative Ophthalmology & Visual Science*, 58(3), 1650-1656.

[10]. Shin, H. C., Roth, H. R., Gao, M., et al. (2016). Deep convolutional neural networks for computer-aided detection: CNNs for CAD. *IEEE Transactions on Medical Imaging*, 35(5), 1285-1298.

[11]. Kermany, D. S., Goldbaum, M., Cai, W., et al. (2018). Identifying medical diagnoses and treatable diseases by image-based deep learning. *Cell*, 172(5), 1122-1131.e9.

[12]. Howard, A. G., Zhu, M., Chen, B., et al. (2017). MobileNets: Efficient convolutional neural networks for mobile vision applications. *arXiv preprint arXiv:1704.04861*.

[13]. Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. In *Advances in neural information processing systems* (pp. 1097-1105). (Note: Cite the actual RFMiD dataset source if available, this is a general ImageNet citation often associated with pre-training). *Self-correction: You should cite the specific RFMiD dataset paper or source.* Example: Pachade, S., et al. (2021). Retinal Fundus Multi-disease Image Dataset (RFMiD): A dataset for multi-disease detection research. *Data in Brief*, 35, 106823.