

# AI-Powered Disease Detection and Diagnosis for Mango Cultivation

<sup>1</sup>Fouzia Anjum S, <sup>2</sup>Dr Arunalatha J S

<sup>12</sup>Department of Computer Science and Engineering

<sup>12</sup>University of Visvesvaraya College of Engineering, Bangalore, India

<sup>1</sup>[fouziaanjum2001@gmail.com](mailto:fouziaanjum2001@gmail.com), <sup>2</sup>[arunalathajs@uvce.ac.in](mailto:arunalathajs@uvce.ac.in)

**Abstract**—Fruit crops such as Mango is widely cultivated in India which accounts for about 45% of the Global Production of Mangoes. It is one of the major nutrient source which is an essential part of a human's diet. However, it is also prone to a lot of diseases which needs to be timely attended. Early Recognition of the diseases of Mango leaves is essential for prevention of further deterioration of the crop. This work focuses on an AI-driven approach for identifying the Mango Leaf Diseases by utilizing FasterViT Model and Retrieval Augmented Generation (RAG) approach for intelligent contextual evaluation. It overcomes the hallucinations issue of the LLMs and leverages the disease identification and Treatment suggestions. The designed system integrates the Classification model with the language models to overcome the limitations corresponding with the Large Language Models (LLMs). The suggested model aims to develop sustainable Agricultural Practices by promoting Precision Agriculture. The proposed model alleviates the dependency on Pesticides by early Evaluation and Remediation of the diseases of Leaves of Mango Trees. Various Deep Learning models viz., ResNet18, VGG16, DenseNet121 etc., are discussed here to generate the insights. The experimental Results demonstrate a remarkable accuracy of 99.88% over the other Benchmark Models [17][22][30].

**Index Terms**— Deep Learning, FasterViT, LLM, Mixtral 8x7b instruct, RAG, Sustainable Agriculture

## I. INTRODUCTION

Agriculture is undoubtedly the largest livelihood contributor in India. Rural areas account for a major contribution in the Agricultural sector. Cultivation of crops is mediated by various attributes namely quality of soil, water, temperature, humidity etc. Farmers are wholly dependent on Agriculture for their subsistence. They cultivate variety of crops such as Fruits, Vegetables, and Grains. It is the backbone of every country and is the most basic amenity for human survival. One such fruit crop Mango which is identified as “King of Fruits” is significantly cultivated worldwide especially in parts of Asia. It is a prominent horticultural crop which is harvested in India and is a major contributor for the country's economy [1]. India occupied a top place in 2010 with the yield of 37.12 million tons and a total area of 2,312 thousand hectares which accounts for around 40% of world's gross mango yield [2]. With more than a thousand different types, India produces the largest varieties of Mangoes [3]. It is a rich source of antioxidants having phytochemicals such as mangiferin and gallotannins [4]. In spite of such a worldwide recognition of this particular fruit crop, it faces a significant challenge due to the variety of Diseases it is prone to. The detection of diseases remains a challenge due to its diverse range and limited availability. Farmers have to bear significant loss of capital, time and resources when dealing with it. It is also difficult for Phytopathologists to accurately detect the diseases with naked eye. Mango leaves are also widely used in rituals and has medicinal properties. Its fruit is known to be rich in Vitamins A and C [5]. It is liable to various diseases in particular Powdery mildew, Anthracnose, Red rust, Bacterial black spot, Ganoderma root rot, lichens, Golmachi, Mango malformation, shuti mold, moriche disease etc. which have negative impact in the production [6] [7]. These diseases not only hamper the growth of the crop but also pose a substantial risk in the economic success at a global level. Furthermore, many diseased leaves resemble each other which pose a challenge to accurately identify the specific disease which is also time consuming, therefore an automated approach would benefit the farmers and phytopathologists [8]. Artificial Intelligence is fundamental in the detection of the diseases. Various Machine learning models namely K-means, SVM, KNN, Random Forest and others have been utilized earlier for the classification of the diseases on various datasets. The authors in [9] give a review on both Machine Learning and Deep Learning Models for determining diseases of plants. Deep Learning shows more promising results in Computer Vision tasks including Image Classification.

## A. Motivation

Agriculture is a paramount for a nation's overall development. There is a dire need to address the issue of sustainable agriculture to overcome the disruption of shortages of food. The following research directs to address the disparity amidst the Agricultural Sector and Technology by focusing on the shortcomings between them. The motive at the core of this research is to provide an efficient framework for the classification of mango diseases and give real time treatment suggestions for the early management of the crops. Proposed approach is to utilize a hybrid model of both Convolutional Neural Networks and Vision Transformers which have exhibited better performance in the classification tasks. Prompt detection will result in an enhanced production of crops and will help the farmers to act quickly [10]. The remediations will ensure timely treatments of the diseases. Cutting edge technologies which are on rise in this changing environment hold promise for sustainable agriculture.

## B. Contribution

The major contributions of this work are summarized below:

- A Dataset named MangoLeafBD [11] having 4000 images containing 8 Classes of diseases specifically Anthracnose, Cutting Weevil, Die Back, Bacterial Canker, Gall Midge, Sooty Mould, Powdery Mildew and Healthy is utilized as shown in Fig. 1.
- Various Data Preprocessing techniques are applied viz., Resizing, Normalizing and Flipping the images. Optimizer and Loss Function parameters are also used.

- Transfer Learning models namely ResNet18, VGG16, MobileNetV2, DenseNet121, EfficientNetB0, ViT and FasterViT are utilized and compared.
- Evaluation measures including Accuracy, Loss, Recall, Precision and F1-score are calculated and Confusion Matrix is illustrated .
- RAG-LLM approach is incorporated for generating context-aware treatment options.
- Finally, a Gradio based Web Application is created for a user-friendly interface.



Fig. 1: MangoLeafBD [11] Dataset representing 8 classes containing Diseased and Healthy Images

## II. LITERATURE SURVEY

Various experiments have been undergone by Researchers to develop effective solutions for the classification of the Mango Leaf Diseases. Machine Learning & Deep Learning Frameworks have been experimented. Mia *et al.*, [12] utilized machine learning based SVM achieving a classification score of 80%, they have introduced Neural network Ensemble model which helps to effortlessly identify the indicators of Mango Leaf Diseases. Ritika *et al.*, [13] proposed a Machine Learning approach wherein they created a python based web application which is helpful for farmers in pesticide application and pest control, they achieved an accuracy of 43% with the help of SVM. Later they also implemented CatBoost and XGBoost models. They achieved an accuracy of 72.05% with the help of a tailored CNN model. Saleem *et al.*, [14], introduced a unique segmentation approach in order to segment the affected region by reviewing patterns of the vein of mango leaves. Feature extraction and fusion was done by the analysis of canonical correlation (CCA) and classification using cubic SVM achieving an accuracy of 95.5%. Pham *et al.*, [15] developed a neural network for early identification of mango leaf disease, which is helpful for low-spec devices such as smartphones. They achieved an accuracy of 77.16% but it fails to fine-tune the selected features such as hidden node count, layer count and activation function. Swapno *et al.*, [16] presented a novel machine learning for assessing mango leaf diseases using EfficientNetv2L architecture achieving a testing accuracy score of 94.99% on the MangoLeafBD Dataset. Rabbi *et al.*, [17] have created a hybrid model by integrating Deep Learning models Inceptionv3 & DenseNet121 attaining 97.92% accuracy. Banerjee *et al.*, [18] aims to diagnose powdery mildew disease of mango leaf based on four acuity levels by making use of a hybrid SVM-CNN approach. It achieves an accuracy of 89.29%. The model gives a false negative ratio which demonstrates a potential need for improvement in characterizing similar acuity levels. Kumar *et al.*, [19] proposed a novel approach of classifying the powdery mildew severity levels of mango leaves by using a combination of CNN & Random Forest. Rizvee *et al.*, [20] presented a custom CNN model achieving 98.55% accuracy. The model is also designed to reduce the computational complexity. Miah *et al.*, [21] developed a specialized model of EfficientNet with channel attention approach, giving an accuracy of 98.88%. An explainable AI method was used to interpret the proficiency of the model. Prabu *et al.*, [22] utilized a pre-trained MobileNetv2 model with a cross over levy flight distribution algorithm for optimizing the hyperparameters and better feature selection, also Augmentation techniques were applied to prevent over-fitting, finally Support Vector Machine is used at the end for the identification of disease. The advocated model obtains an accuracy of 94.5%, recall of 98.03%, 99.56% of precision and an f1-score of 95.67%. Gautam *et al.*, [23] presented an aggregated deep learning model which is designed to solve the problem of automatic identification of diseases of the leaves of Mango. The images are segmented and fed as input to a series of deep neural networks which is further integrated with Machine Learning model to characterize the disease. Seetha *et al.*, [24] proposed a novel Hybrid Coyote Grey Wolf Optimization (CO-GWO) algorithm. The images are segmented and then the classification is done by the neural network classifier and the weights are adjusted by the CO-GWO algorithm which enhances the classification accuracy. It achieves an accuracy of 96.7111%. Wang *et al.*, [25] proposed a lightweight YOLOv8 model, GAS-YOLOv8 which decreased the model parameters by 33% while improving the accuracy, the mean average precision for pests of Mango was around 98% and the mean average precision of diseases was around 91%. Admass *et al.*, [26] implemented a hybrid Convolutional Neural Networks & Histogram Oriented Gradients (CNN-HOG) Approach wherein the CNN was used for feature extraction while the HOG was used for acquiring the shape & Texture data. The proposed model achieved accuracy of 99.5%. Selvakumar *et al.*, [27] proposed a systematized model for the classification of leaf diseases of Mango using an optimized Fuzzy C Means (FCM) based segmentation, Deviation based updated Dingo Optimizer (D-UDOX), Weighted Optimized Recurrent Neural Network (WO-RNN) and ResNet-150 achieving an accuracy of 96%. Gowrishankar *et al.*, [28] implemented ConvNeXt Models for the Mango Leaf Disease diagnosis, ConvNeXtLarge Model has showed highest accuracy of 98.79% for the category of mango pests, mango pathogens have achieved 100% and combined dataset is of 99.17%. The larger models have outperformed the smaller models. Mahbub *et al.*, [29] designed

a unique Lightweight Convolutional Neural Network (LCNN) to classify the MangoLeafBD Dataset which achieves highest testing accuracy of 98%, various baseline models such as ResNet50, VGG16, Xception have also been tested, and is found that LCNN surpasses these models pertaining to accuracy. Bezabh *et al.*, [30] proposed a stacking of VGG16 & GoogLeNet models of CNN, various image preprocessing techniques viz., image resizing, noise reduction, and image augmentation, noise removal operations, segmentation techniques viz., K- Means and Mask-RCNN, data augmentation procedures and feature extraction operations were applied. Testing accuracy of 99.21% was obtained. Singh *et al.*, [8] developed a Deep Transfer Learning Driven (DTLD) model to characterize the diseases of Mango leaves. Various algorithms such as DenseNet, MobileNet *etc.*, dataset and validation methods have been trained and tested on the model. Softmax activation function is applied to classify the disease, the model which is developed has obtained an accuracy score of 99.76%.

### III. PROBLEM FORMULATION

#### Problem Statement:

To develop an Automated AI based system using a hybrid Convolutional Neural Network-Transformer and an LLM- RAG approach for the detection of Mango Leaf Diseases.

#### The Objectives are to:

- To improve the Accuracy and Reliability.
- Generate context aware response for various unpredictable factors such as soil conditions, climate conditions and the evolving disease-causing organisms.

### IV. PROPOSED METHODOLOGY

This section describes the methodology of the proposed system wherein seven different Deep learning Models are analyzed which are used for the purpose of Mango Disease Classification. The various phases involved in the proposed system are described subsequently. Phase-1 is the Data Collection, Phase-2 describes the Data Pre-processing and Augmentation, Phase-3 involves Training the 7 different models, in Phase-4 we test the models, Phase-5 involves Prompt generation, Phase-6 is the RAG-LLM framework, Phase-7 is the creation of User Interface for integrating the Trained model with RAG- LLM.

#### A. Data Collection

An open-source Dataset from Mendeley named MangoLeafBD [11] is used. It contains 4000 images of 8 classes, with each class containing 500 images. The names of the classes of disease are Anthracnose, Powdery Mildew, Bacterial Canker, Sooty Mould, Gall Midge, Die back, Cutting Weevil respectively and there is also a Healthy Class. The images are of size 240x320 pixels having RGB colour. Of these 4000 images, 1800 are of distinct leaves. These images were acquired from various Mango Trees via the camera of a mobile phone as shown in Fig. 1.

#### B. Data Pre-processing and Data Augmentation

One of the key aspects of Data Augmentation is to reduce the overfitting nature of a model. It is also required to ensure that the dataset is Balanced and robust in nature. In the MangoLeafBD [11] Dataset Data Augmentation has been carried out by Zooming, Rotating, Image Resizing, Noise Cleaning thereby increasing the count of the images to 4000. This technique is useful when the Dataset is small.

Rotation technique is used to rotate the image by a certain angle, Zooming is done to randomly generate images by zooming it in or out, Resizing, shearing, translation, flipping are the commonly used Geometric transformations [31].

In Deep Learning, Data pre-processing such as Normalize, Centre crop, Resize, Horizontal crop is performed by various Python frameworks such as Pytorch, Keras, Tensorflow which have in-built functions for image transformations [31].

The Dataset is Stratified into Training and Testing Set. The Training Set is used for calibrating the model, whereas the Testing Set is used for evaluation with sample images.

#### C. Model Training

This section aims to study seven different models that are utilized to train the classification task of mango leaf disease. Transfer Learning approach is followed by utilizing the Pre-Trained Deep Learning Architectures and fine tuning it for the purpose of classification of the Mango Disease.

#### MODEL 1: RESNET18

It is a type of Convolutional Neural Network architecture. ResNet18[32] has 18 layers, 17 being convolutional layers and the terminal layer is a fully connected layer. This pretrained model was trained on a subset of ImageNet database. Here the model is fine-tuned by freezing the convolutional layers and modifying the classification layer for the task of multi class classification. The pre-trained feature extraction capabilities are retained for the mango leaf disease classification. The Residual block contains a shortcut connection which adds the input  $x$  to the output of Stacked layers  $F(x)$  as shown in Equation (1).

$$F(x) = H(x) + x \quad (1)$$

where:

- $x$  is input to the residual block
- $F(x)$  is output of residual block
- $H(x)$  is output of Stacked layers

**MODEL 2: VGG16**

VGG16 [33] is a variant of Convolutional neural network that has thirteen convolutional layers, five pooling layers and three fully connected layers. This pretrained model was trained on a subset of ImageNet database. Here the model is fine-tuned by freezing the convolutional layers and modifying the classification layer for the task of multi class classification. The pre-trained feature extraction capabilities are retained for the mango leaf disease classification.

**MODEL 3: DENSENET121**

The DenseNet [34] is a model from the Densely connected Convolutional Networks that integrates each neural layer to all the other neural layers in a feed-forward arrangement. It has five convolution and pooling layers, one classification layer, three transition layers and two dense blocks. This pretrained model was trained on a subset of ImageNet database. Here the model is fine-tuned by freezing the convolutional layers and modifying the classification layer for the task of multi class classification. The pre-trained feature extraction capabilities are retained for the mango leaf disease classification.

**MODEL 4: EFFICIENTNETB0**

EfficientNetB0 is a variant of Convolutional Neural Network which is trained on ImageNet Database. The EfficientNet Model [35] which is proposed, introduced a new scaling method which uniformly scales all dimensions of width(W), depth(D) and resolution(R) using a compound coefficient. The scaling formula is given by Equation (2).

$$W * D^2 * R^2 \approx Constant \quad (2)$$

Here the model is pre-trained and adapted by locking up all the convolutional layers & modifying the classification layer for the task of multi class classification. The pre-trained feature extraction capabilities are retained for the mango leaf disease classification.

**MODEL 5: MOBILENETV2**

MobileNetV2 is another variant of Convolutional Neural Network which is trained on ImageNet Database that uses inverted residual blocks with bottlenecking features [36] as given in Equation (3).

$$Output = Proj(DWConv(Exp(Input))) \quad (3)$$

where:

- *Input* is the input feature maps
- *Proj* is the Projection layer
- *DWConv* is the Depthwise convolution
- *Exp* is the expansion layer

Here the model is pre-trained and adapted by locking up all the convolutional layers & modifying the classification layer for the task of multi class classification. The pre-trained feature extraction capabilities are retained for the mango leaf disease classification.

**MODEL 6: ViT**

Vision Transformer is a variant of Deep Learning architecture. This architecture was originally proposed for the scope of natural language processing (NLP) wherein it appeared to be first mentioned in [37] and presented in [38]. It divides the image into a series of regions and each region is then converted into a vector representation, the vector embeddings are then fed into a Transformer encoder. The model can then make predictions by using the classification token in the series of patch embeddings. Both local and global representations are captured due to the mechanism of self-attention of ViT. Unlike CNN, ViT captures the long-range dependencies within an image [37].

**MODEL 7: FASTERViT**

Fast Vision Transformers with Hierarchical Attention (FasterViT) [39] is a hybrid CNN-ViT based architecture which was developed by researchers at NVIDIA, it was trained on the ImageNet Database. It provides high performance and requires reduced computational resources for tasks like image classification, object detection. Due to inadequacy of Inductive Bias in Vision Transformers, it requires a lot of training data and may hinder the performance. FasterViT resolves this issue without requiring large training dataset. FasterViT is based on the proposed Hierarchical Attention.

The Hierarchical Attention Mechanism(HAT) mechanism is specifically designed to handle high resolution images. It disassociates global self-attention mechanism into multi-level attention mechanism. It makes use of carrier tokens which interact over the entire feature map and facilitate efficient cross window communication. The HAT uses window based self-attention and makes use of the above carrier tokens. The outputs of local self-attention is combined with global self-attention to retrieve the long range dependencies within the images.

The HAT uses a window based self-attention which limits the attention computation to smaller regions of the input. The computational complexity grows linear. The overview of FasterViT Architecture is discussed subsequently. FasterViT consists of 4 different stages. Each stage uses different types of blocks. First an input image is fed into the network. Stage 1 and 2 uses Convolutional Blocks for local Feature Extraction while the Stage 3 and Stage 4 uses the Hierarchical attention blocks for capturing long range. Each stage is followed by a downsampling block that minimizes the granularity of feature maps. Stage 1 and 2 blocks are responsible for local feature extraction from the input images, each convolutional block have convolutional layers, activation function and batch normalization. Stage 3 and 4 blocks capture long range dependencies. Some convolutional operations are replaced with attention mechanisms to understand complex spatial relationship from the input. The FasterViT model mentioned above for the classification of mango Leaf Diseases is used. Fig. 2 represents the Proposed Architecture of the model. Adam Optimizer is utilised for the reduction of loss during the process of Training and the Cross-Entropy is used as a Loss function to generate an ideal solution.

---

**Algorithm 1** Prediction of Mango Leaves Diseases and Treatment Suggestion Pipeline

---

- 1: **Input:** MangoLeafBD Dataset
  - 2: **Output:** Predict Disease Type and provide Treatment Suggestions
  - 3: Import the MangoLeafBD Dataset from Drive
  - 4: Import necessary libraries: torch, matplotlib, torchvision, etc.
  - 5: Get pretrained weights for FasterViT. Setup the FasterViT model instance with pretrained weights and Freeze the base parameters
  - 6: Get automatic transforms from pretrained FasterViT weights, Turn the image into Data Loaders
  - 7: Create optimizer and loss function
  - 8: Train the classifier head of the Pre-Trained FasterViT feature extractor Model
  - 9: Install the necessary requirements and upload the Document Source
  - 10: Create the RAG chain: Load the document and split it into chunks, Create embeddings and index them using FAISS
  - 11: Generate the Disease specific query and integrate with LLM to generate context-aware responses
  - 12: Deploy the system using Gradio and Upload an image to predict disease type and treatment
  - 13: Plot the Accuracy, Loss Curves, and other Performance Metrics
- 

#### D. Model Testing

To test the images the Dataset is stratified into both Training and Testing Set. The models are evaluated on this test set to generate the Evaluation metrics namely Accuracy, Loss curves, Precision, F1-score and Recall. The plot of the confusion matrix is to find the classification proficiency of each model. It compares the predicted values to actual values.

Accuracy is determined as the total number of inferences that were genuinely correct. It is represented as in Equation (4)

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FN + FP} \quad (4)$$

Loss is the error between the inference values and the target values.

Precision is the number of positive predictions that were genuinely found correct. It is represented as in Equation (5)

$$\text{Precision} = \frac{TP}{TP + FP} \quad (5)$$

Recall is the actual favourable cases that the model duly identified as positive. It is represented as in Equation (6)

$$\text{Recall} = \frac{TP}{TP + FN} \quad (6)$$

where :

TP: True Positive

TN: True Negative

FP: False Positive

FN: False Negative

F1-Score is the combined score of Recall & Precision and it is represented as in Equation (7)

$$\text{F1-Score} = \frac{2 * \text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \quad (7)$$

The models were trained and tested in a Google Colaboratory Setup which provides free access to Computational Resources such as GPUs and TPUs. Each model is trained separately for 10 epochs, and a T4 GPU is utilized which is provided under free-tier access by Google Colaboratory and record the results as shown from Fig. 3 to Fig. 17.

Fig. 3 describes the Accuracy Metric of all the models for a Dataset split of Ratio 80:20, as can be seen the FasterViT achieves an Accuracy of 99.88%. Fig. 4 describes the Loss curves for all the models. The FasterViT model gives the least Loss of 0.93%. Fig. 5 gives the Accuracy Metric of all the models for the ratio of 70:30, here too the FasterViT model achieves 99.75%. Fig. 6 gives the Loss metrics for all the models. Fig. 7 gives the Accuracy metric for a split ratio of 60:40, here both ViT and FasterViT models give an accuracy of 99.63%. Fig. 8 describes the Loss metrics for 60:40, FasterViT model gives a loss of 2%. Fig. 9 shows the Accuracy curve of models for 50:50 split ratio and Fig. 10 gives the Loss metrics for the same split. As can be seen the Accuracy increases as we increase the ratio of Training Data to Testing Data. Loss value decreases by increasing the Training data ratio.

The ResNet18 Model achieves a Testing Accuracy of 98.5% & Training Accuracy of 98.62%. Testing Loss of 2.33% & Training Loss of 4.64% respectively. Table 1 depicts the Evaluation metrics for each class. Fig. 11 represents the Confusion Matrix for ResNet18 model.

The VGG16 model attains a Testing Accuracy of 98.75% & Training Accuracy of 97.31%. Testing Loss of 1.042% and Training Loss of 1.396% respectively. Table 2 represents the Evaluation metrics for each class. Fig. 12 represents the Confusion Matrix for VGG16 model.

The DenseNet121 Model attains a Testing Accuracy of 99.12% & Training Accuracy of 98%. Validation loss of 2.23% and Training Loss of 2.894% respectively. Table 3 represents the Evaluation metrics for each class. Fig. 13 represents the Confusion Matrix for DenseNet121 model.

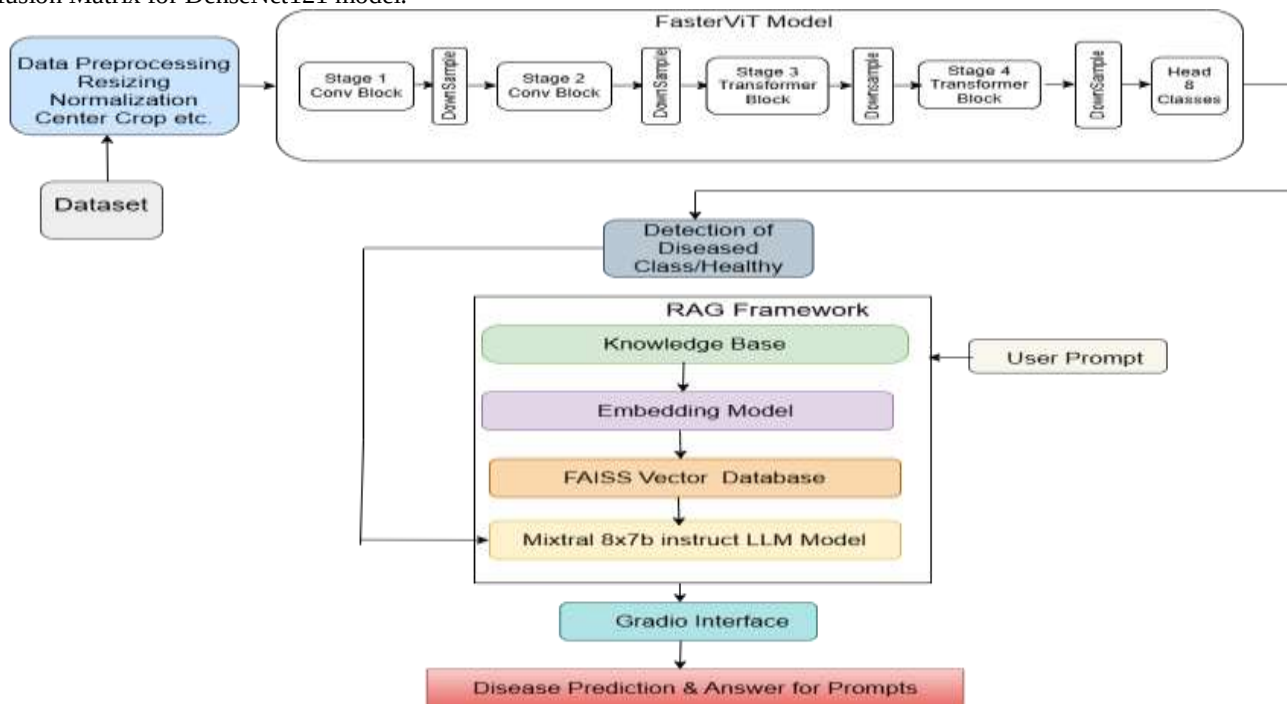


Fig. 2: Workflow of Proposed Methodology

TABLE I: Precision, Recall and F1-score for ResNet18 Model

| Class Names      | Precision | Recall | F1-Score |
|------------------|-----------|--------|----------|
| Anthracnose      | 0.9298    | 1.0000 | 0.9636   |
| Bacterial Canker | 1.0000    | 1.0000 | 1.0000   |
| Cutting Weevil   | 1.0000    | 1.0000 | 1.0000   |
| Die Back         | 1.0000    | 0.9683 | 0.9839   |
| Gall Midge       | 1.0000    | 1.0000 | 1.0000   |
| Healthy          | 1.0000    | 1.0000 | 1.0000   |
| Powdery Mildew   | 1.0000    | 0.9811 | 0.9905   |
| Sooty Mould      | 1.0000    | 0.9773 | 0.9885   |

TABLE II: Precision, Recall and F1-score for VGG16 Model

| Class Names      | Precision | Recall | F1-Score |
|------------------|-----------|--------|----------|
| Anthracnose      | 1.0000    | 0.9700 | 0.9800   |
| Bacterial Canker | 1.0000    | 1.0000 | 1.0000   |
| Cutting Weevil   | 1.0000    | 1.0000 | 1.0000   |
| Die Back         | 1.0000    | 1.0000 | 1.0000   |
| Gall Midge       | 0.9700    | 0.9800 | 0.9800   |
| Healthy          | 0.9600    | 1.0000 | 0.9800   |
| Powdery Mildew   | 0.9800    | 0.9900 | 0.9800   |
| Sooty Mould      | 0.9900    | 0.9600 | 0.9700   |

**TABLE III: Precision, Recall and F1-score for DenseNet121 Model**

| Class Names      | Precision | Recall | F1-Score |
|------------------|-----------|--------|----------|
| Anthracnose      | 1.0000    | 1.0000 | 1.0000   |
| Bacterial Canker | 1.0000    | 1.0000 | 1.0000   |
| Cutting Weevil   | 1.0000    | 1.0000 | 1.0000   |
| Die Back         | 1.0000    | 1.0000 | 1.0000   |
| Gall Midge       | 0.9700    | 0.9800 | 0.9900   |
| Healthy          | 0.9800    | 1.0000 | 0.9900   |
| Powdery Mildew   | 1.0000    | 1.0000 | 1.0000   |
| Sooty Mould      | 0.9800    | 1.0000 | 0.9900   |

**TABLE IV: Precision, Recall and F1-score for EfficientNetB0 Model**

| Class Names      | Precision | Recall | F1-Score |
|------------------|-----------|--------|----------|
| Anthracnose      | 0.9800    | 1.0000 | 0.9900   |
| Bacterial Canker | 1.0000    | 1.0000 | 1.0000   |
| Cutting Weevil   | 1.0000    | 1.0000 | 1.0000   |
| Die Back         | 1.0000    | 1.0000 | 1.0000   |
| Gall Midge       | 1.0000    | 1.0000 | 0.9900   |
| Healthy          | 1.0000    | 1.0000 | 1.0000   |
| Powdery Mildew   | 0.9800    | 0.9800 | 0.9800   |
| Sooty Mould      | 0.9800    | 0.9800 | 0.9800   |

**TABLE V: Precision, Recall and F1-score for MobileNetV2 Model**

| Class Names      | Precision | Recall | F1-Score |
|------------------|-----------|--------|----------|
| Anthracnose      | 0.9800    | 1.0000 | 0.9900   |
| Bacterial Canker | 1.0000    | 1.0000 | 1.0000   |
| Cutting Weevil   | 1.0000    | 1.0000 | 1.0000   |
| Die Back         | 0.9800    | 1.0000 | 0.9900   |
| Gall Midge       | 1.0000    | 0.9800 | 0.9900   |
| Healthy          | 0.9800    | 1.0000 | 0.9900   |
| Powdery Mildew   | 1.0000    | 1.0000 | 1.0000   |
| Sooty Mould      | 1.0000    | 1.0000 | 1.0000   |

**TABLE VI: Precision, Recall and F1-score for ViT Model**

| Class Names      | Precision | Recall | F1-Score |
|------------------|-----------|--------|----------|
| Anthracnose      | 1.0000    | 1.0000 | 1.0000   |
| Bacterial Canker | 1.0000    | 1.0000 | 0.9900   |
| Cutting Weevil   | 0.9900    | 1.0000 | 1.0000   |
| Die Back         | 1.0000    | 1.0000 | 1.0000   |
| Gall Midge       | 1.0000    | 0.9900 | 1.0000   |
| Healthy          | 1.0000    | 1.0000 | 1.0000   |
| Powdery Mildew   | 0.9900    | 1.0000 | 0.9900   |
| Sooty Mould      | 1.0000    | 0.9900 | 0.9900   |

**TABLE VII: Precision, Recall and F1-score for FasterViT Model**

| Class Names      | Precision | Recall | F1-Score |
|------------------|-----------|--------|----------|
| Anthracnose      | 1.0000    | 1.0000 | 1.0000   |
| Bacterial Canker | 1.0000    | 1.0000 | 1.0000   |
| Cutting Weevil   | 1.0000    | 1.0000 | 1.0000   |
| Die Back         | 1.0000    | 1.0000 | 1.0000   |
| Gall Midge       | 1.0000    | 1.0000 | 1.0000   |
| Healthy          | 1.0000    | 1.0000 | 1.0000   |
| Powdery Mildew   | 0.9900    | 1.0000 | 1.0000   |
| Sooty Mould      | 1.0000    | 0.9900 | 0.9900   |

The EfficientNetB0 Model attains a Testing Accuracy of 99.25% & Training Accuracy of 99.25%. Testing loss of 2.39% and Training Loss of 4.54% respectively. Table 4 represents the Evaluation metrics for each class. Fig. 14 represents the Confusion Matrix for EfficientNetB0 model.

The MobileNetV2 Model attains a Testing Accuracy of 99.75% and Training Accuracy of 99.03%. Testing Loss of 1.07% and Training Loss of 3.03% respectively. Table 5 represents the Evaluation metrics for each class. Fig. 15 represents the Confusion Matrix for MobileNetV2 model. The ViT Model attains a Testing Accuracy of 99.62% & Training Accuracy of 100%. Testing Loss of 0.13% and Training Loss of 0.66% respectively. Table 6 represents the Evaluation metrics for each class. Fig. 16 represents the Confusion Matrix for ViT model.

The FasterViT Model obtains a Testing accuracy of 99.88% & Training Accuracy of 97.63%. Testing Loss of 0.93% & Training Loss of 8.61% respectively. Table 7 represents the Evaluation metrics for each class. Fig. 17 represents the Confusion Matrix for FasterViT model.

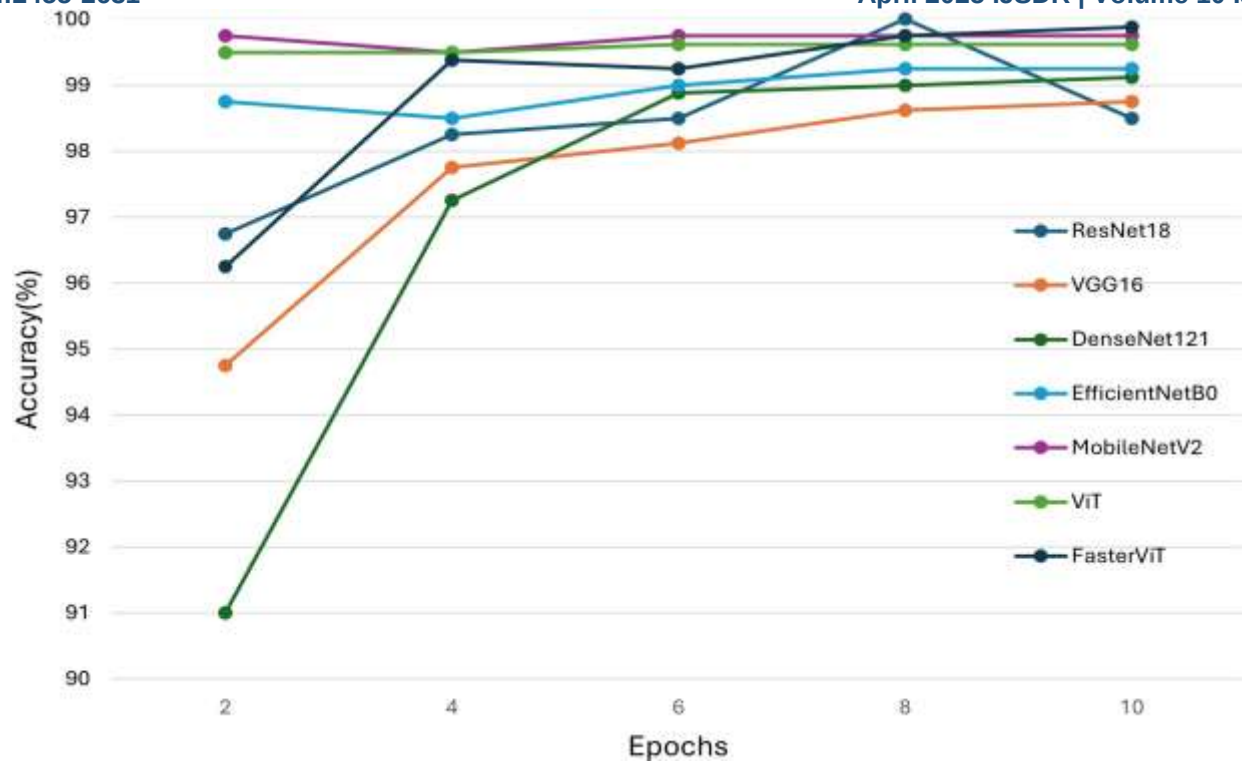


Fig. 3: Accuracy for 80:20 Dataset MangoLeafBD[11]

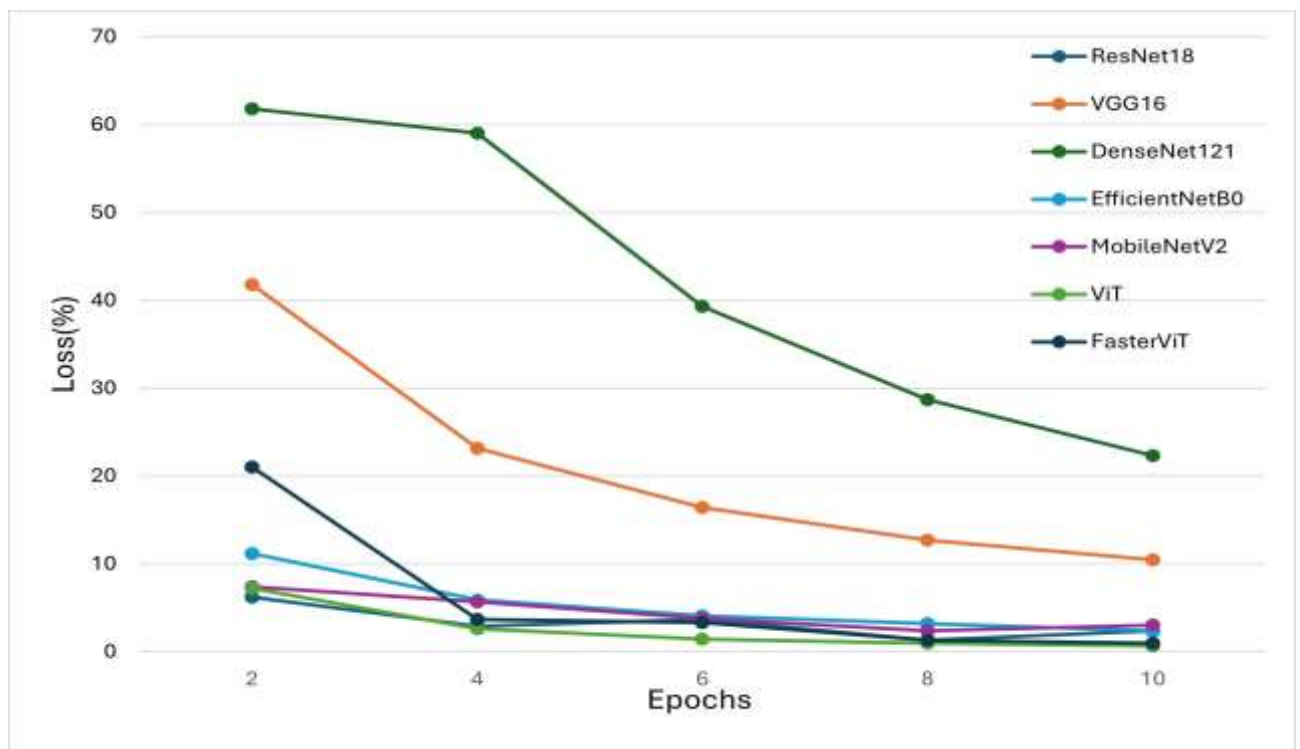


Fig. 4: Loss for 80:20 Dataset MangoLeafBD[11]

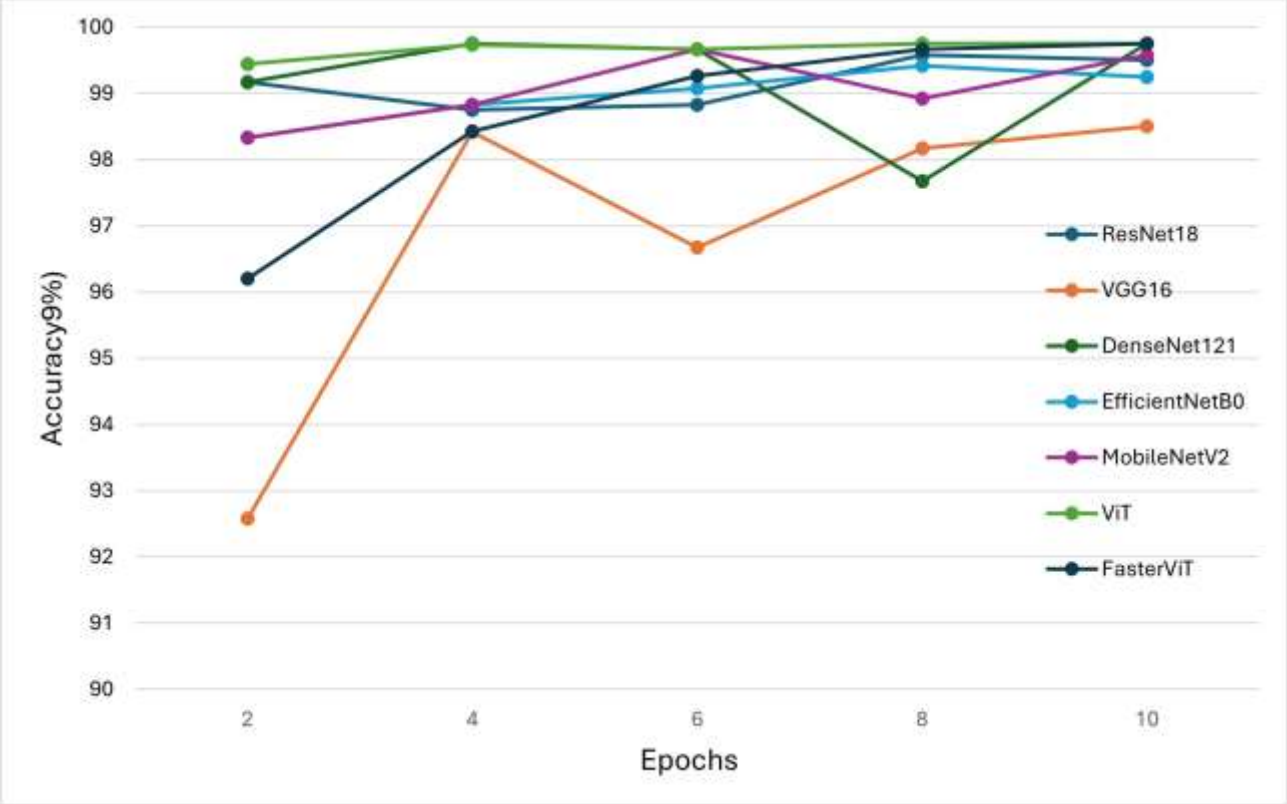


Fig. 5: Accuracy for 70:30 MangoLeafBD[11] Dataset

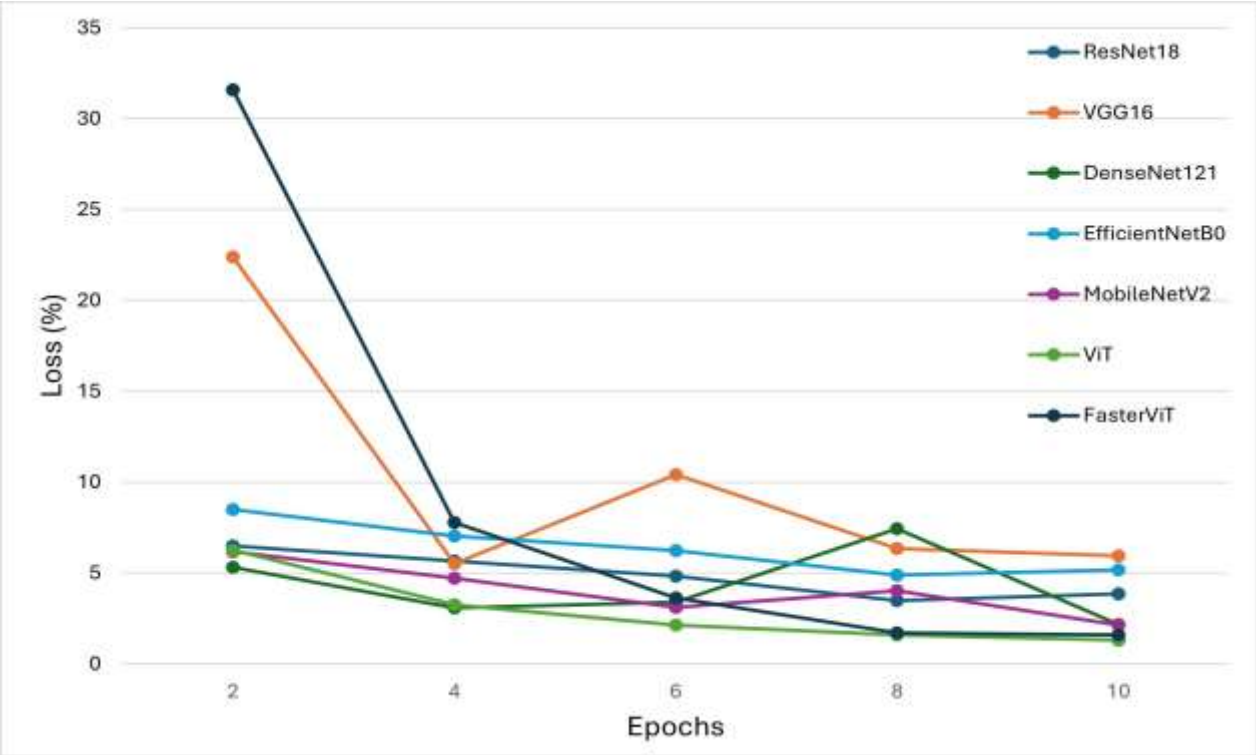


Fig. 6: Loss for 70:30 MangoLeafBD[11] Dataset

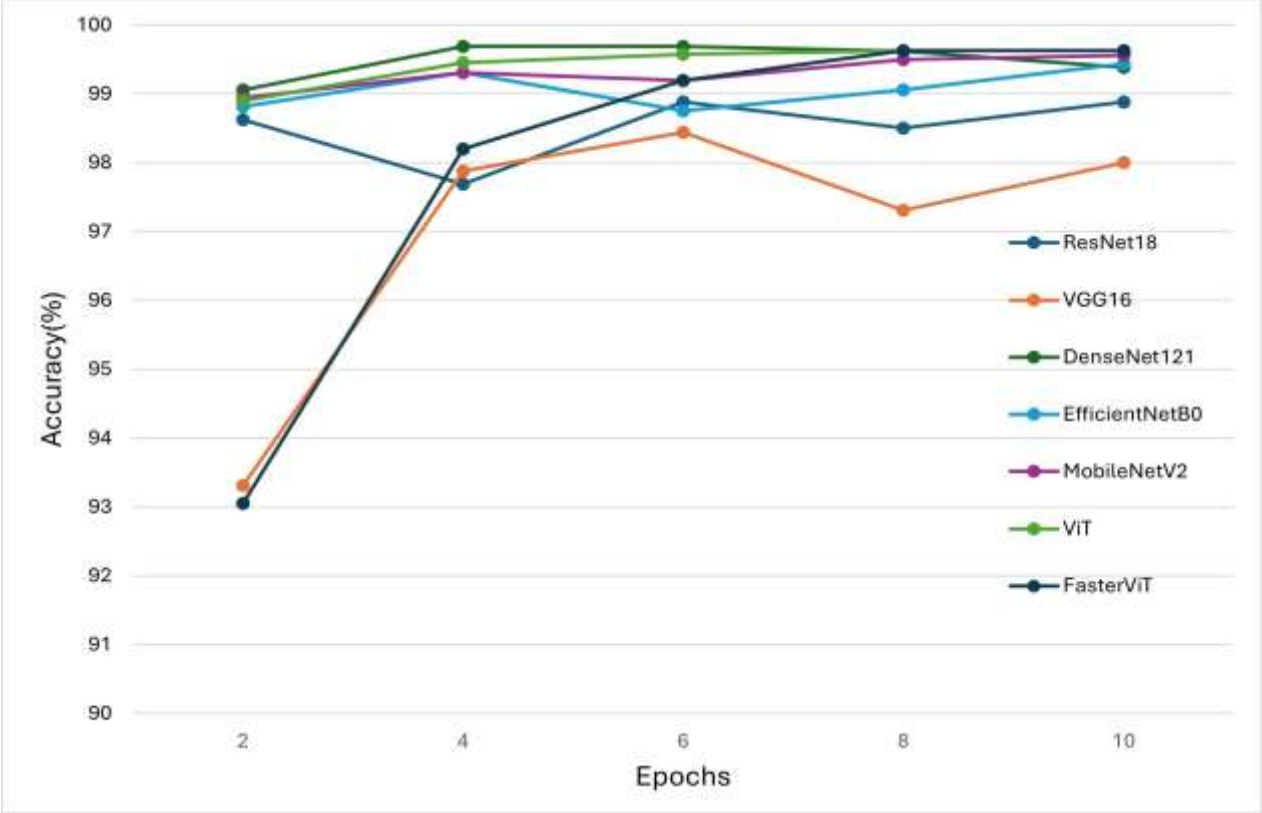


Fig. 7: Accuracy for 60:40 MangoLeafBD[11] Dataset

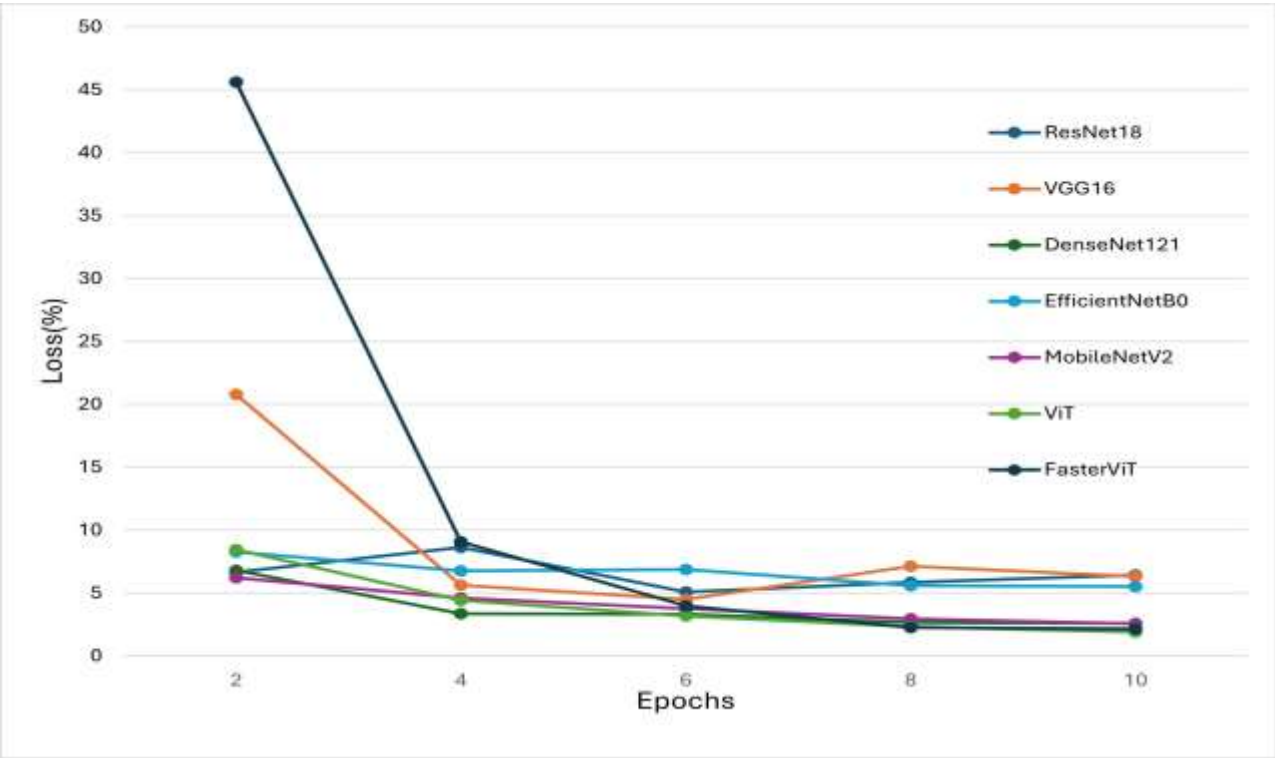


Fig. 8: Loss for 60:40 MangoLeafBD[11] Dataset

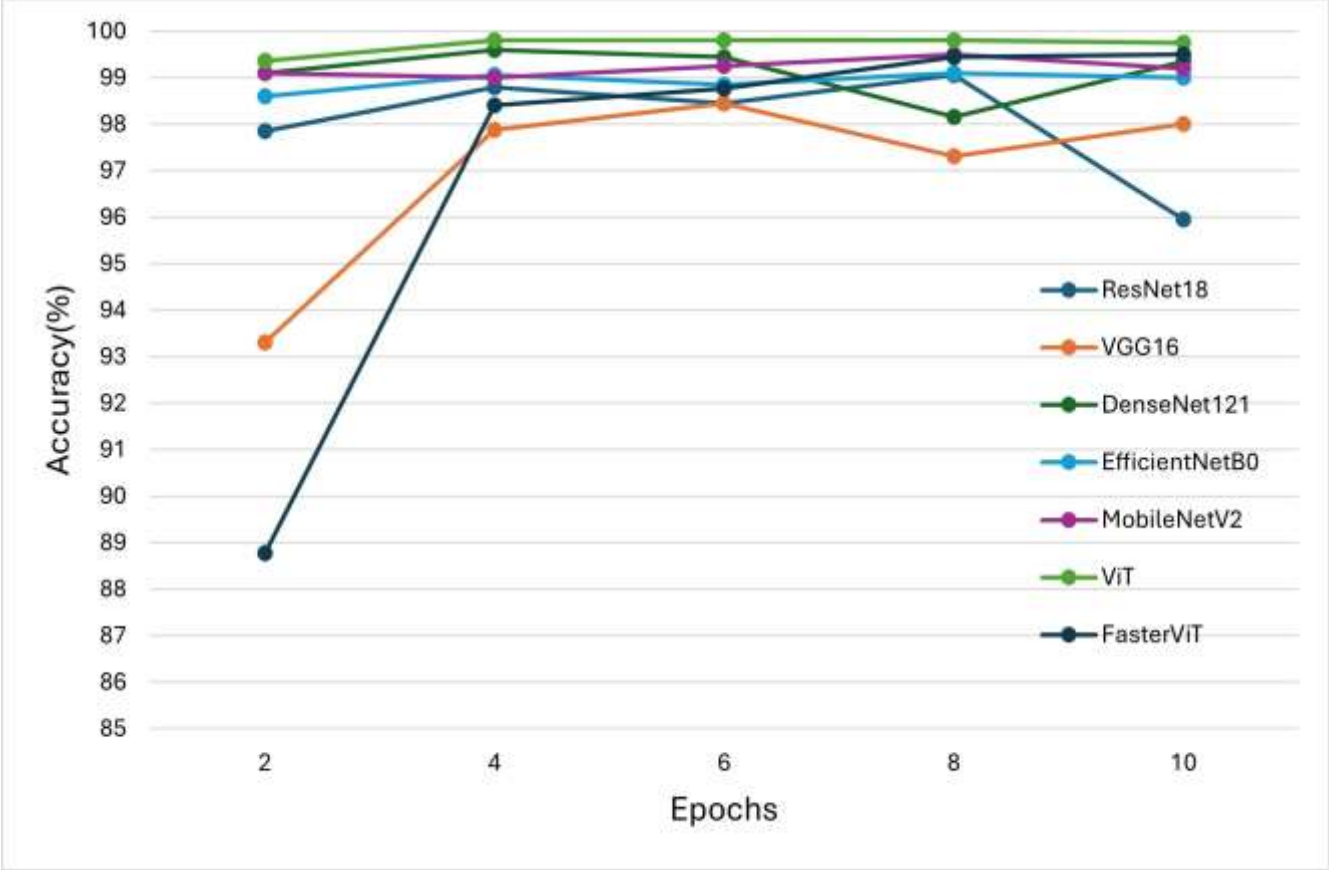


Fig. 9: Accuracy for 50:50 MangoLeafBD[11] Dataset

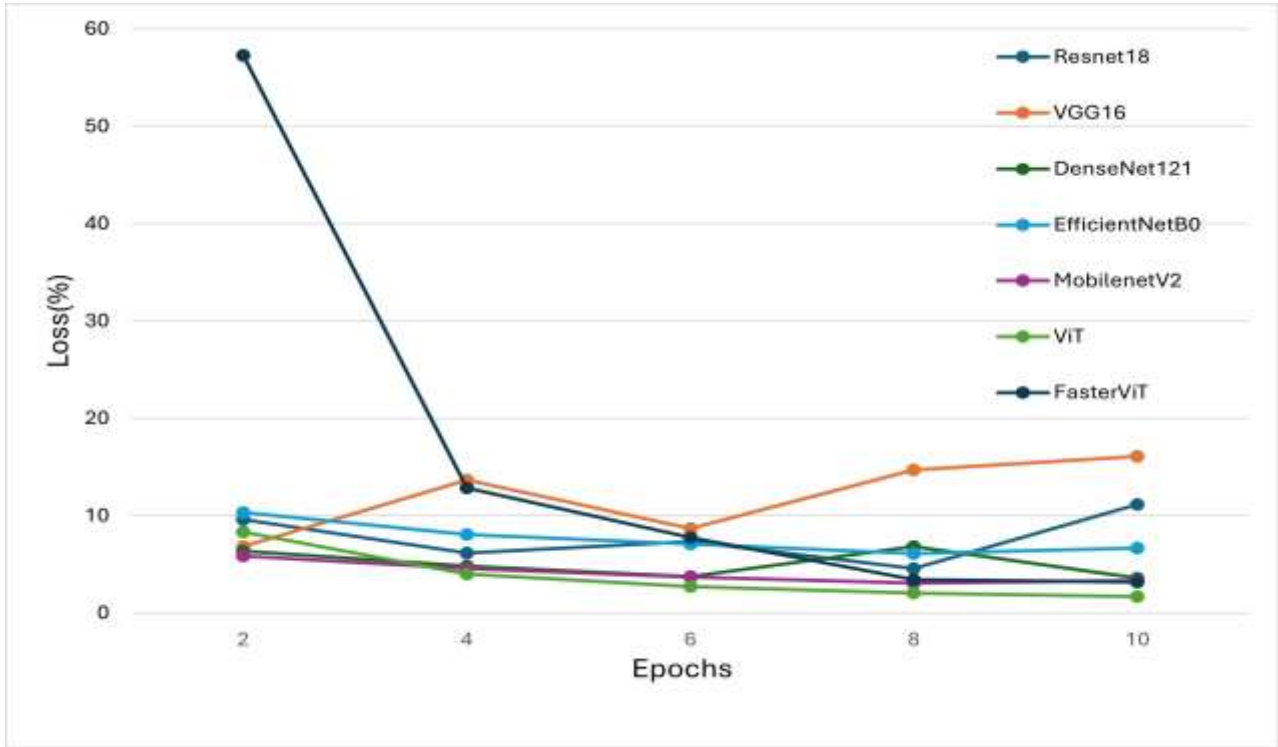


Fig. 10: Loss for 50:50 MangoLeafBD[11] Dataset

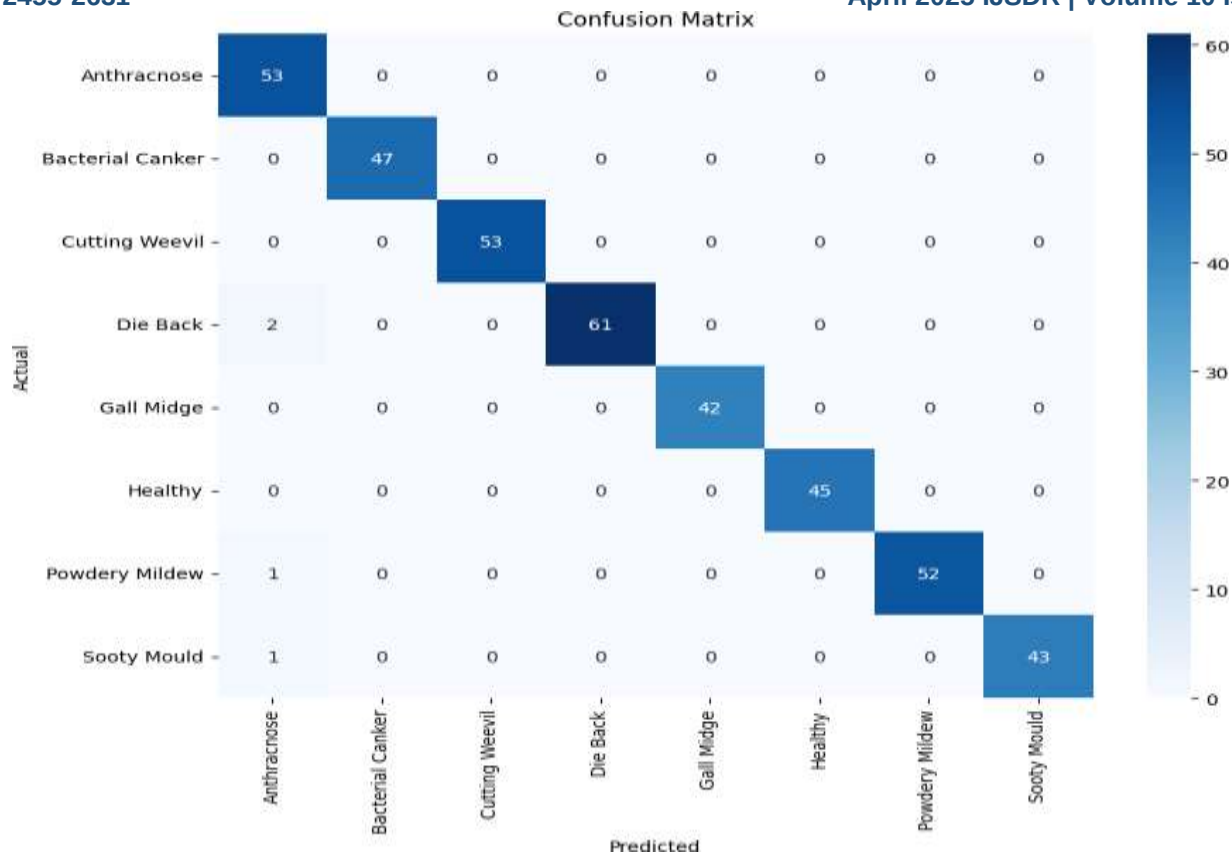


Fig. 11: Confusion Matrix of ResNet18 Model

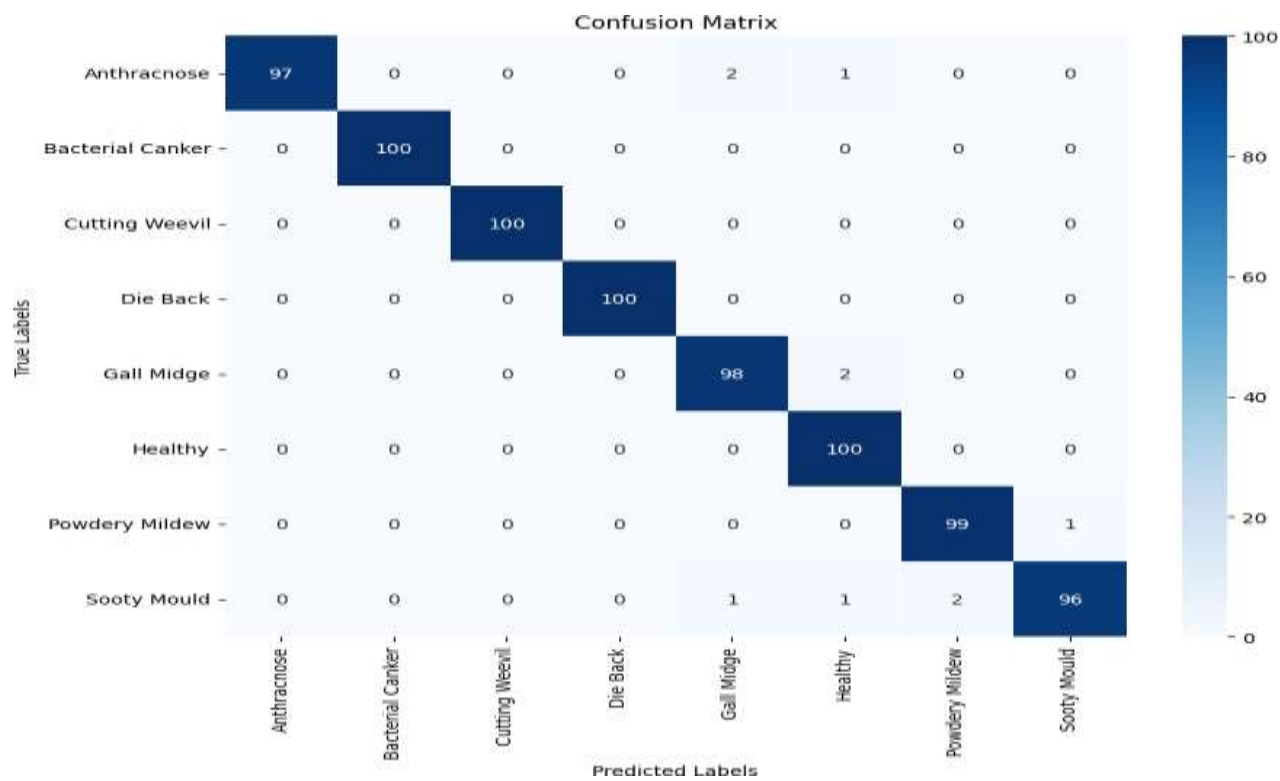


Fig. 12: Confusion Matrix of VGG16 Model

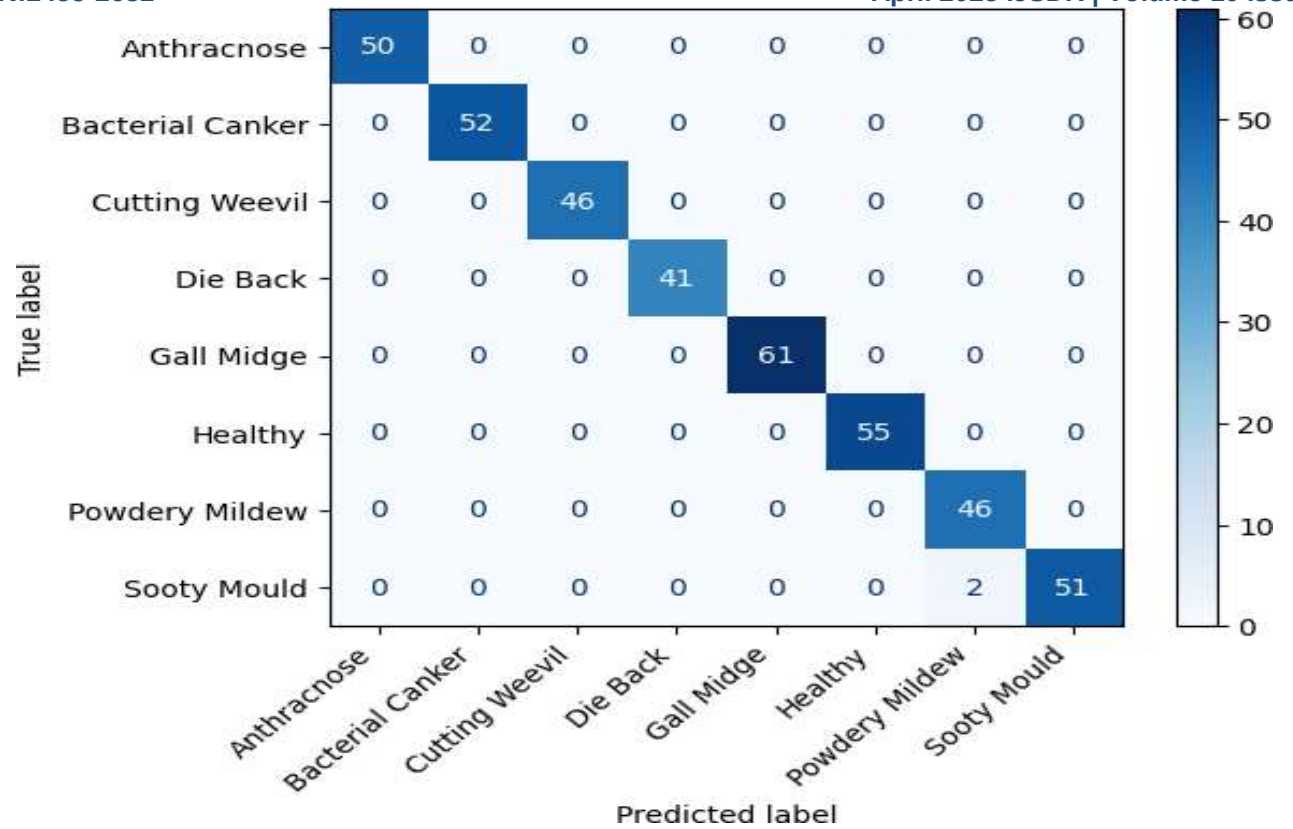


Fig. 13: Confusion Matrix of DenseNet121 Model

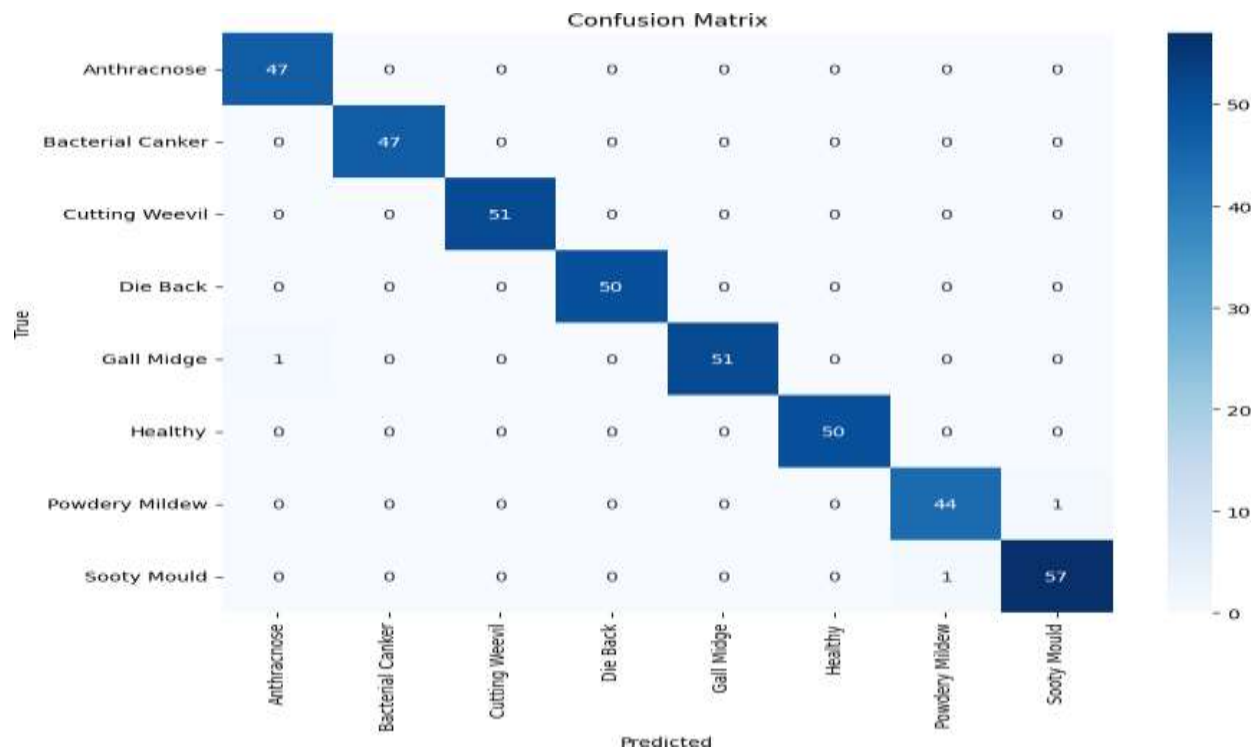


Fig. 14: Confusion Matrix of EfficientNetB0 Model

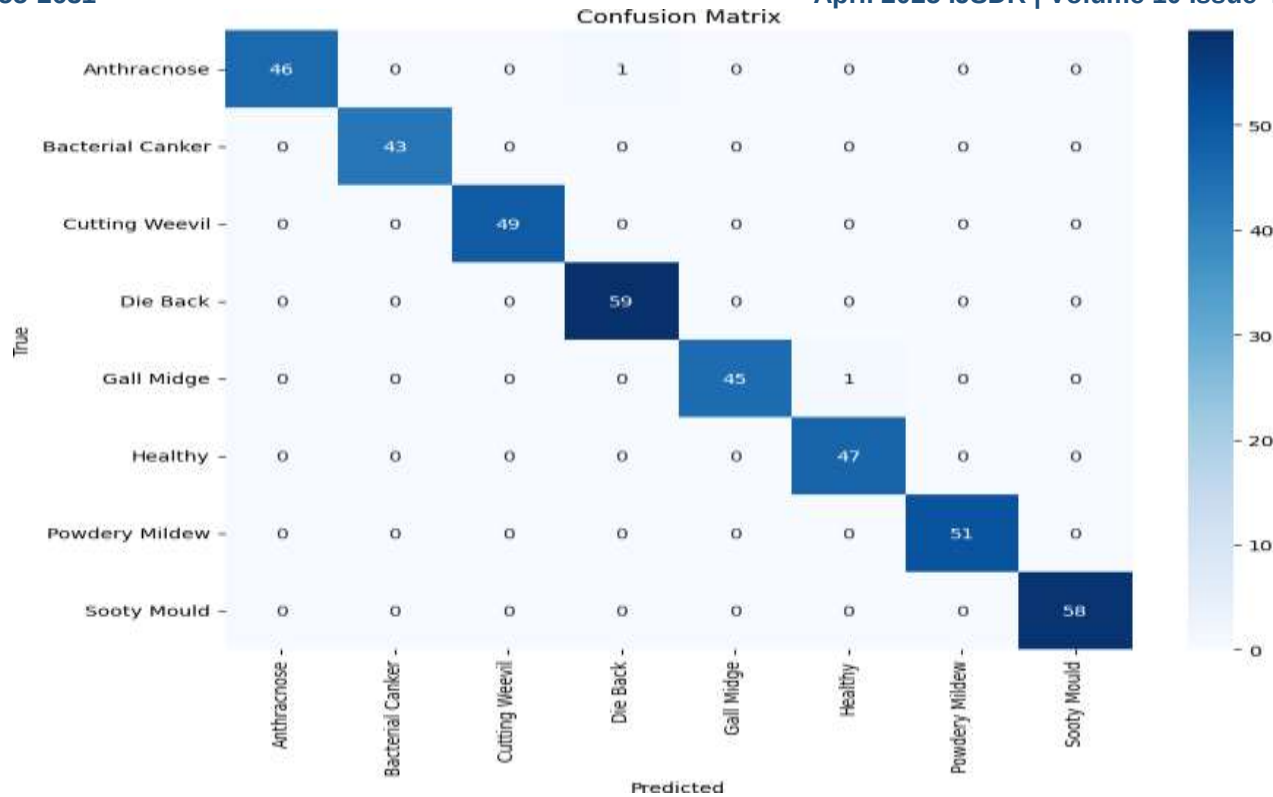


Fig. 15: Confusion Matrix of MobileNetV2 Model

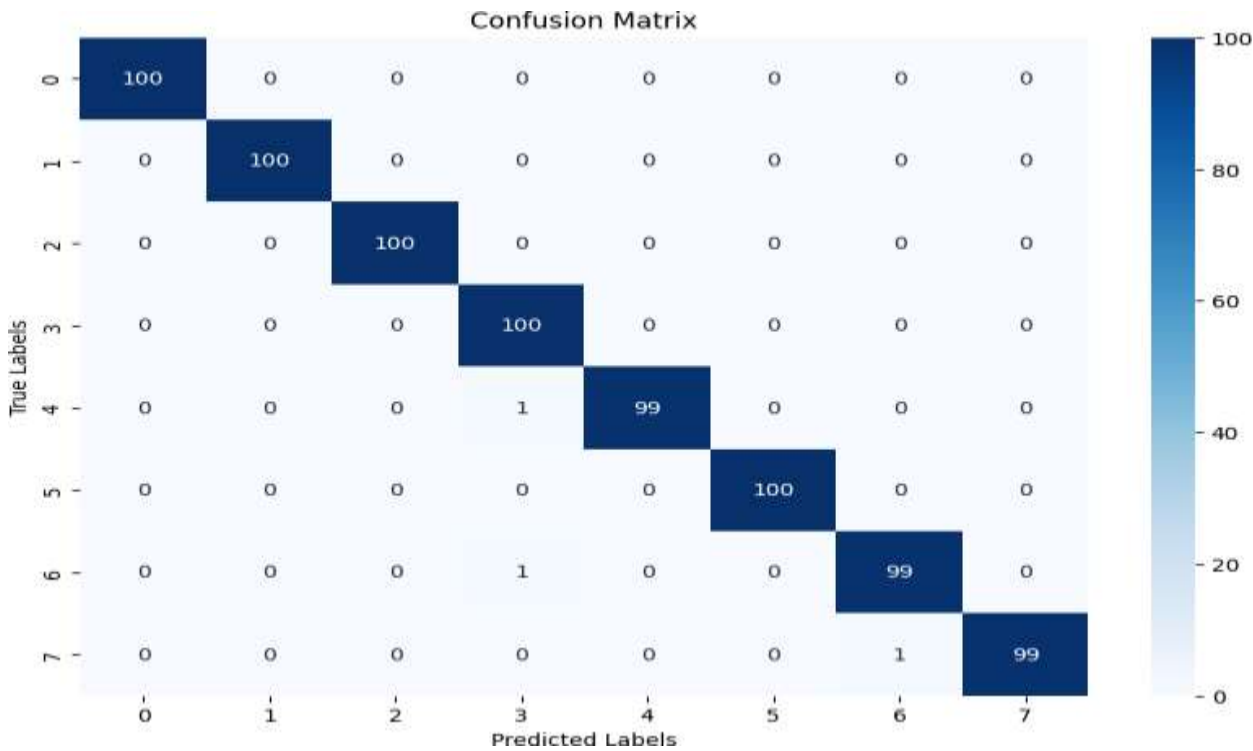


Fig. 16: Confusion Matrix of ViT Model

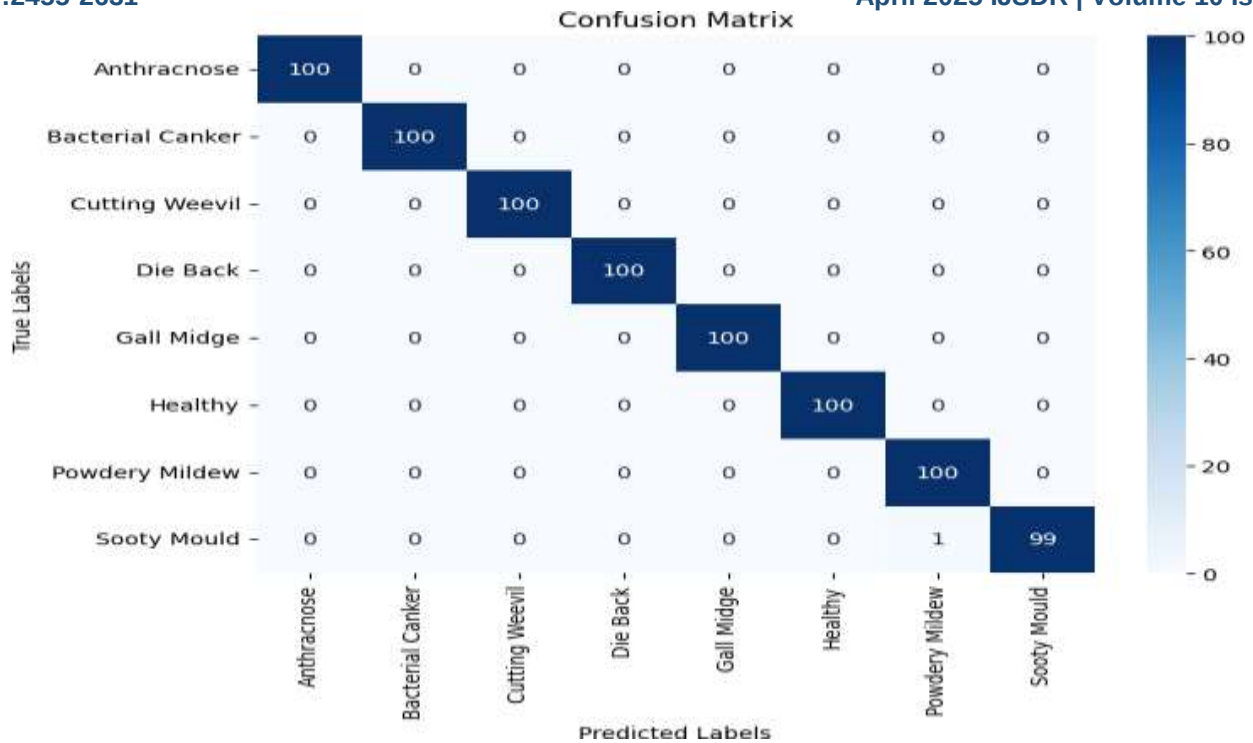


Fig. 17: Confusion Matrix of FasterViT Model

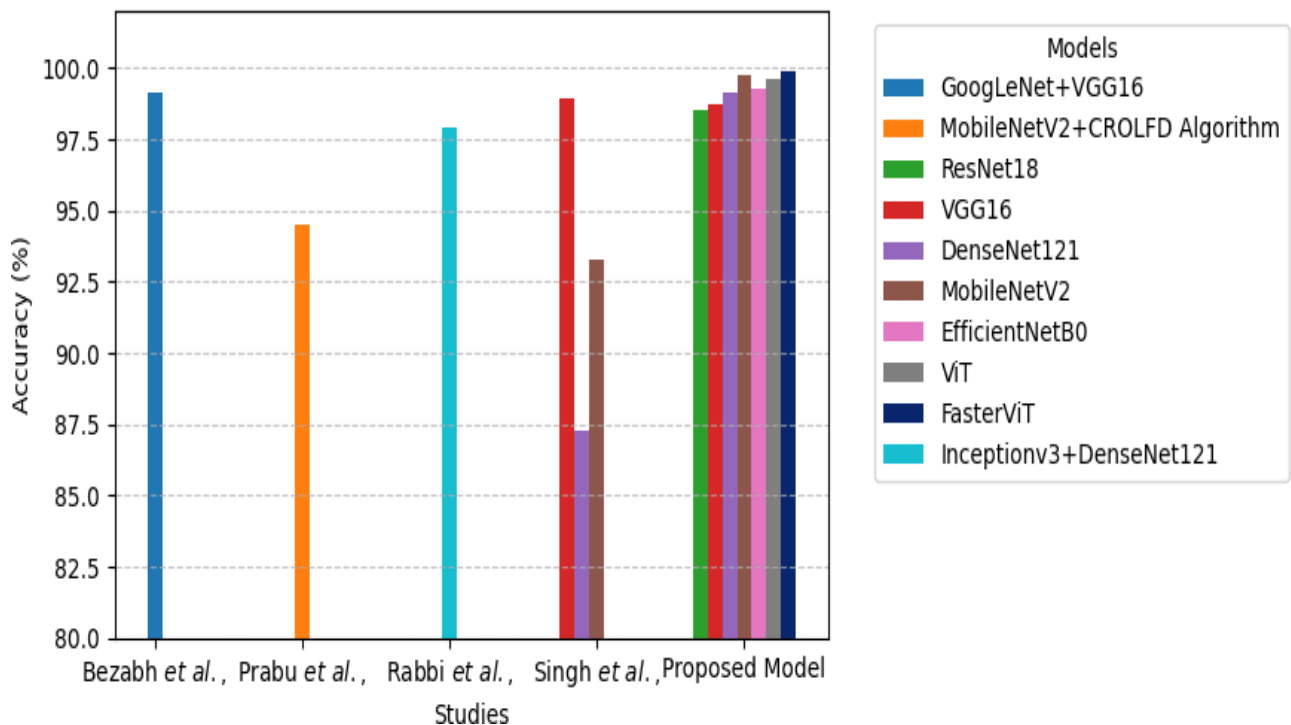


Fig. 18: Comparative analysis of Proposed Work with Existing work

### E. Prompt Generation

In order to generate desired outcome and detailed responses from the Language models it is necessary to specify the context-specific, clear prompts which act as queries to the LLM's. The generated prompts must be dataset specific with precise information about the mango leaf diseases. This will help in generating fine-tuned outputs and save time.

### F. RAG – LLM Framework

Once the FasterViT model is trained, the predicted class label is passed to the RAG-LLM framework as shown in Fig. 2. RAG [40] which is known as Retrieval Augmented generation helps to generate accurate and reliable information by retrieving information from external sources which helps the LLMs in generating real time context-specific responses. This helps the problem

of Hallucination by LLMs. RAG model is built using Langchain framework which easily helps to integrate it with the LLM. A Knowledge Base containing information about each class of disease is developed, which contains textual information and is segmented into manageable chunks. These chunks are embedded using the Hugging Face embedding for semantic representation. The embeddings are then indexed using a Facebook AI Similarity Search (FAISS) which contribute to the creation of Vector store enabling similarity based retrieval. A RAG chain is then created for user friendly interaction to gather relevant information. Mixtral 8x7b instruct LLM model is used which enables the Language model to generate relevant real time responses by integrating the document information and the LLM understanding. This leverages the AI driven capabilities, thereby efficiently diagnosing the Mango leaf disease in the Real-time and suggesting the treatment remedies. Retrieving document information proves to be more efficient in this scenario.

### G. User Interface

After the development of RAG-LLM framework an open- source python package “Gradio” is used which allows to quickly build a web interface for the Model. The user can upload an image of a mango leaf disease, the interface responds with the type of disease and also suggests remedies for that particular disease, user can also manually type query related to plant health.

## III. RESULT AND ANALYSIS

This section gives the Results wherein the comparison of various Deep Learning models have been done. Analysis of existing work is also performed. Fig. 18 represents the comparison of various Existing Models. As illustrated from the Fig. 18, FasterViT surpasses other Deep learning Models by giving an impressive accuracy of 99.88%.

Fig. 18 gives insights on the existing work done. As can be seen Prabu *et al.*, [22] have utilized a Pre-Trained MobileNetV2 model along with CROFLD Algorithm and have attained an accuracy of 94.5%. Bezabh *et al.*, [30] developed a hybrid VGG16 & GoogLeNet model based upon CNN, various Segmentation Techniques viz., K-Means and Mask- RCNN, Image Preprocessing Methods, Noise Removal Strategies, Data Augmentation Approaches and Feature Extraction Operations were applied. The model has acquired a Testing Accuracy of 99.21%. Rabbi *et al.*, [17] developed a hybrid model by integrating Deep Learning models Inceptionv3 and DenseNet121 achieving an accuracy of 97.92%. Therefore the FasterViT based RAG-LLM Model which utilizes both the CNN and Transformer approach of Deep Learning is proposed. This AI based web application will help the Farmers in easy detection of the Mango Leaf Disease and suggest the context-aware remedies in real time. This will help Farmers and Phytopathologists to easily interact with the user-friendly interface.

Table 8 gives the details of Training and Testing Accuracy and Loss for each of the 7 Deep Learning Models. Faster-ViT [39] model gives the highest Testing Accuracy among all the models.

**TABLE VIII: Performance comparison of different models for Mango Leaf Disease Classification**

| Model          | Training Accuracy (%) | Testing Accuracy (%) | Training Loss | Testing Loss |
|----------------|-----------------------|----------------------|---------------|--------------|
| ResNet18       | 98.62                 | 98.50                | 4.64          | 2.33         |
| VGG16          | 97.31                 | 98.75                | 1.396         | 1.042        |
| DenseNet121    | 98.00                 | 99.12                | 2.894         | 2.23         |
| EfficientNetB0 | 99.25                 | 99.25                | 4.54          | 2.39         |
| MobileNetV2    | 99.03                 | 99.75                | 3.03          | 1.07         |
| ViT            | 99.89                 | 99.62                | 0.66          | 0.13         |
| FasterViT      | 97.63                 | 99.88                | 8.61          | 0.93         |

## IV CONCLUSION

The proposed AI-Powered Disease Detection and Diagnosis of Mango Leaves, leveraging the FasterViT model for disease detection and RAG-LLM integration for the context-aware diagnosis is a revolution in the Agriculture Industry. By integrating the classification model with the language model proposed model addresses the challenges posed in the prompt diagnosis and Remediation of Mango Leaf Diseases. While classification of the diseases using the FasterViT model is performed, the challenge lies with the LLM’s Hallucinating. Therefore, by introducing the Retrieval Augmented Generation (RAG) the model gives context-specific diagnosis by using the external data source. This enhances the Accuracy and Reliability of the overall system leading to Precision Agriculture. This helps in the reduction of the use of Pesticides and promotes eco-friendly agricultural practices. It also brings a positive change in the agriculture industry by promoting Reliability and Scalability of the crop production. The future scope of this work would be to include more classes of disease and various kinds of food crops viz., Rice, Wheat, Millets *etc.*, This system would bring a positive change in the agricultural industry by increasing the yields of the crop, securing the food supplies and promoting the economic success at a global level. In this transformative field of agriculture, the proposed AI- powered system introduces a revolutionary change, shaping the future of agricultural technology.

## REFERENCES

- [1] Chakma S, "Adapting Land Degradation and Enhancing Ethnic Livelihood Security Through Fruit Production: Evidence from Hilly Areas of Bangladesh," *Agro-biodiversity and Agri-ecosystem Management*, Springer, Singapore, pp. 217–238, 2022.
- [2] Pradeep Kumar, A.K Misra and Dinesh Modi, "Current Status of Mango Malformation in India," *Asian Journal of Plant Sciences*, vol. 10, pp. 1-23, 2011.
- [3] Lockwood R, "The Mango. Botany, Production and Uses," *Experimental Agriculture*, vol. 46, no. 3, pp. 416-416, 2010.
- [4] S Naik and P Desai, "Mango (*Mangifera indica* L.) Classification Using Convolutional Neural Network and Linear Classifiers," in *Proceedings of the Third International Conference on Sustainable Computing. Advances in Intelligent Systems and Computing*, vol. 404, Springer, Singapore, 2022.
- [5] Wasyihun Sema, Yirga Yayeh and Girmaw Bogale, "Automatic Detection and Classification of Mango Disease Using Convolutional Neural Network and Histogram Oriented Gradients," 2023.(Preprint)
- [6] D Sabareeswaran and R. Guna Sundari, "A Hybrid of Plant Leaf Disease and Soil Moisture Prediction in Agriculture using Data Mining Techniques," *International Journal of Applied Engineering Research*, vol. 12, no. 18, pp. 7169-7175, 2017.
- [7] Richard Strange, Peter Scott, "Plant Disease: A Threat to Global Food Security", *Annual Review of Phytopathology*, no. 43. pp. 83-116, 2005.
- [8] Y.P Singh, B.K. Chaurasia and M.M.Shukla, "Deep Transfer Learning Driven Model for Mango Leaf Disease Detection," *International Journal of System Assurance Engineering and Management*, vol. 15, pp. 4779–4805, 2024.
- [9] Jackulin C and Murugavalli S, "A Comprehensive Review on Detection of Plant Disease using Machine Learning and Deep Learning Approaches," *Journal of Measurement: Sensors*, vol. 24, 2022.
- [10] Tawseef Shaikh, Tabasum Rasool and Faisal Lone, "Towards Leveraging the Role of Machine Learning and Artificial Intelligence in Precision Agriculture and Smart Farming," *Journal of Computers and Electronics in Agriculture*, vol. 198, pp. 107-119, 2022.
- [11] Sarder Iftekhar Ahmed, Muhammad Ibrahim, Md. Nadim, Maria Mehjabin Shejunti, Taskeed Jabid and Md. Sawkat Ali, "MangoLeafBD: A Comprehensive Image Dataset to Classify Diseased and Healthy Mango Leaves," *Journal of Data in Brief*, vol. 47, ISSN 2352-3409, 2023.
- [12] M.R. Mia, S. Roy and S.K. Das "Mango Leaf Disease Recognition using Neural Network and Support Vector Machine," *Iran Journal of Computer Science* vol. 3, pp. 185–193, 2020.
- [13] Ritika Pramod Chendvenkar, "Identification and Classification of Leaf Pests within the Indonesian Mango Farms using Machine Learning", 2021.
- [14] R. Saleem, J.H. Shah, M. Sharif, M. Yasmin, H.S. Yong and J. Cha, "Mango Leaf Disease Recognition and Classification Using Novel Segmentation and Vein Pattern Technique," *Applied. Sciences*, vol. 11, no. 24, 2021.
- [15] Tan Pham, Ly Tran and Son Dao, "Early Disease Classification of Mango Leaves Using Feed-Forward Neural Network and Hybrid Metaheuristic Feature Selection," *IEEE Access*, vol. 8, pp. 189960 – 189973, 2020.
- [16] Masfequier S M, "A Novel Machine Learning Approach for Fast and Efficient Detection of Mango Leaf Diseases," in *Proceedings of the IEEE Third International Conference on Computing and Machine Intelligence (ICMI)*, no. 1-7, 2024.
- [17] Munshi Rabbi, Israt Jahan, Lamia Rukhsara and Tapasy Rabeya, "Deep Learning for Detection of Mango Leaf Disease: A Comparative Study Using Convolutional Neural Networks Models," in *Proceedings of the Sixth International Conference on Electrical Engineering and Information & Communication Technology (ICEEICT)*, Dhaka, Bangladesh, pp. 675-680, 2024.
- [18] Deepak Banerjee, Vinay Kukreja, Shanmugasundaram Hariharan and Vishal Jain, "Enhancing Mango Fruit Disease Severity Assessment with CNN and SVM-Based Classification," in *Proceedings of the IEEE Eight International Conference for Convergence in Technology (I2CT)*, Lonavla, India, pp. 1-6, 2023.
- [19] Sandeep Kumar and Bhupesh Gupta, "Hybrid CNN-Random Forest Framework for Elevated Accuracy in Powdery Mildew Disease Classification," in *Proceedings of the Third International Conference for Innovation in Technology (INOCON)*, Bangalore, India, pp. 1-6, 2024.
- [20] Redwan Rizvee, Tasnim Orpa, Adil Ahnaf, Md Kabir, Mohammad Rashid, Mohammad Islam, Maheen Islam, Taskeed Jabid and Md Ali, "LeafNet: A Proficient Convolutional Neural Network for Detecting Seven Prominent Mango Leaf Diseases," *Journal of Agriculture and Food Research*, vol. 14, 2023.
- [21] Parvez Miah, Chamak Saha, Somak Saha and Muhammad Iqbal Hossain, "Optimizing Disease Identification in Mango Leaves: A Channel Attention Approach with Enhanced EfficientNet," in *Proceedings of the Sixth International Conference on Electrical Engineering and Information & Communication Technology*, pp. 1193-1198, 2024.
- [22] M. Prabu and B.J. Chelliah, "Mango Leaf Disease Identification and Classification using a CNN Architecture Optimized by Crossover-based Levy Flight Distribution Algorithm," *Journal of Neural Computing and Applications*, vol. 34, pp. 7311–7324, 2022.
- [23] V. Gautam, R.K. Ranjan and P. Dahiya, "ESDNN: A Novel Ensembled Stack Deep Neural Network for Mango Leaf Disease Classification and Detection," *Journal of Multimedia Tools and Application*, vol. 83, pp. 10989–11015, 2024.
- [24] J. Seetha, R. Ramanathan and V Goyal, "Mango Leaf Disease Classification using Hybrid Coyote-Grey Wolf Optimization tuned Neural Network Model," *Multimedia Tools and Applications*, vol. 83, pp. 17699–17725, 2024.
- [25] J. Wang and J. Wang, "A Lightweight YOLOv8 based on Attention Mechanism for Mango Pest and Disease Detection," *Journal of Real- Time Image Processing*, vol. 21, no. 136, 2024.
- [26] W.S. Admass, Y.Y. Munaye and G.A. Bogale, "Convolutional Neural Networks and Histogram-Oriented Gradients: a Hybrid Approach for Automatic Mango Disease Detection and Classification," *International Journal of Information Technology*, vol. 16, pp. 817–829, 2024.
- [27] A. Selvakumar and A. Balasundaram, "Automated Mango Leaf Infection Classification Using Weighted and Deep Features with Optimized Recurrent Neural Network Concept," *The Imaging Science Journal*, no. 72 (3), pp. 285–303, 2023.

- [28] Asha Rani K P and Gowrishankar S, "ConvNeXt-based Mango Leaf Disease Detection: Differentiating Pathogens and Pests for Improved Accuracy," *International Journal of Advanced Computer Science and Applications (IJACSA)*, no. 14(6), 2023.
- [29] N. I. Mahbub, F. Naznin, M. I. Hasan, S. M. R. Shifat, M. A. Hossain and M. Z. Islam, "Detect Bangladeshi Mango Leaf Diseases Using Lightweight Convolutional Neural Network," in *Proceedings of the International Conference on Electrical, Computer and Communication Engineering (ECCE)*, Chittagong, Bangladesh, pp. 1-6, doi: 10.1109/ECCE57851.2023.10101648 (2023).
- [30] Yohannes Bezabh, Aleka Ayalew, Biniyam Abuhayi, Tensay Demlie, Eshete Awoke and Taye Endeshaw, "Classification of Mango Disease using Ensemble Convolutional Neural Network," *Journal of Smart Agricultural Technology*, vol. 8, 2024.
- [31] Alhassan Mumuni and Fuseini Mumuni, "Data augmentation: A Comprehensive Survey of Modern Approaches," *Journal of Array*, vol. 16, 2022.
- [32] K. He, X. Zhang, S. Ren and J. Sun, "Deep Residual Learning for Image Recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, pp. 770- 778, 2016.
- [33] K. Simonyan, and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," in *Proceedings of the Third International Conference on Learning Representations, Computational and Biological Learning Society*, pp. 1–14, 2015.
- [34] G. Huang, Z. Liu, L. Van Der Maaten and K. Q. Weinberger, "Densely Connected Convolutional Networks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, USA, pp. 2261-2269, 2017.
- [35] M. Tan and Q.V. Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks," in *Proceedings of the Thirty Sixth International Conference on Machine Learning (ICML)*, Long Beach, pp. 6105-6114, 2019.
- [36] Mark Sandler, "MobileNetV2: Inverted Residuals and Linear Bottle- necks," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 4510-4520, 2018.
- [37] Alexey Dosovitskiy, "An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale," in *Proceedings of the International Conference on Learning Representation*, 2020.
- [38] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, "Attention is All You Need," *Journal of Advances in Neural Information Processing Systems*, vol. 30, 2017.
- [39] Ali Hatamizadeh, Greg Heinrich, Hongxu Yin, Andrew Tao, Jose M. Alvarez, Jan Kautz and Pavlo Molchanov, "FasterViT: Fast Vision Transformers with Hierarchical Attention," in *Proceedings of the Twelfth International Conference on Learning Representations*, 2023.
- [40] P. Lewis, E. Perez, A. Piktus, F. Petroni, V. Karpukhin, N. Goyal, H. Kuttler, M. Lewis, W.T. Yih, T. Rocktaschel and S. Riedel, "Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks," *Journal of Advances in Neural Information Processing Systems*, vol. 33, pp. 9459-9474, 2020.