International Journal of Environmental Sciences
ISSN: 2229-7359
Vol. 11 No. 10s, 2025
https://www.theaspd.com/ijes.php

Papaya Fruit and Leaf Ringspot Virus Detection using Convolutional Neural Network (CNN) and Transfer Learning

Putra Sumari[1]*, Mustafa M. Abd Zaid[2], Ahmed Kateb Jumaah Nussairi[3], Shaffiah binti Mohd[1], Ling Gaomian[1], Devinna Win Anak Boniface Emang[1] and Chew Yee Ling[1]
[1]School of Computer Science, University Sains Malaysia, 11800 Peneng, Malaysia
[2]College of Technical Engineering, Islamic University, Najaf, Iraq
[3]Al-Manara College for Medical Sciences, Amarah, Iraq
Correcponding Author: putras@usm.my

Abstract: *This article uses the application of Convolutional Neural Networks (CNN) and transfer learning models to discover Papaya Ringspot Virus (PRSV) in papaya plants, a critical concern for papaya cultivation in Malaysia. The study compares the performance of CNN with three pre-trained models: VGG16, ResNet50, and InceptionV3. Each model employs the Adam optimizer with a learning rate of 0.001. Models were trained and evaluated utilizing a dataset comprising 1,524 images of papaya fruits and leaves affected by PRSV. CNN, trained for 10 epochs without fine-tuning, achieved a final accuracy of 99.22%, starting from an initial accuracy of 81.91%. In contrast, the pre-trained models demonstrated superior performance. InceptionV3 consistently achieved a perfect final accuracy of 100% across various configurations, with initial accuracies ranging from 92.52% to 96.02%. VGG16 also attained 100% accuracy after fine-tuning. ResNet50 showed robust performance, achieving 100% accuracy in all configurations, with initial accuracies improving with extended training periods. The experimental results indicate that pre-trained models significantly outperform the custom CNN in PRSV detection, highlighting the benefits of transfer learning. InceptionV3, with its efficient architecture and fewer parameters than VGG16, demonstrated the highest performance, making it the most suitable model for this task. The results highlight to ability of deep learning models to enhance and discover crop disease and support sustainable agriculture in Malaysia.*

Keywords: Papaya fruit disease, Custom CNN, VGG16, ResNet50, InceptionV3

1. INTRODUCTION

According to preliminary findings discovered by the Food and Agriculture Organization (FAO), plant diseases account for more than a third of the overall yearly loss in agricultural production [1]. Plant-infectious diseases might cause rapid spread and significant losses. Thus, identification and diagnosis of these diseases are very important. Plant disease identification was traditionally done by agricultural professionals, and it called for great knowledge. This was labor-intensive, difficult task prone to errors [2]. Conventional approaches for detecting plant diseases that rely on manual feature extraction are intricate and ineffective. Progress in computer vision and artificial intelligence, especially in the realm of deep learning, provides answers to many challenges across several domains, including agriculture, yielding more precise outcomes than conventional techniques [3].

The fleshy fruit of the Carica papaya plant, often known as a papaya, papaw, or pawpaw, is a major fruit export for Malaysia [4]. Its rapid growth and early maturation allow for quick returns on investment. The PRSV, or papaya ringspot virus, is a positive-sense RNA virus that belongs to the Potyviridae family and the genus Potyvirus. It may infect papayas [5]. In contrast to other viruses affecting papaya, PRSV spreads rapidly poses the possible dispersion scenarios. Since its 1991 discovery in Johor, the destructive PRSV virus has cut down on papaya production in the southern areas, necessitating a shift of cultivation to central and northern areas [6]. One of Malaysia's most important economic sectors, agriculture accounts for around 8.9 percent of the country's GDP [7]. Crop diseases significantly impact crop yield and quality, highlighting the need for effective disease detection methods [8]. Vitamins C, A, magnesium, folate, fiber, and antioxidants are abundant in papaya [9]. Additionally, papaya's versatility extends to high-end cuisine to proteolytic enzyme, papain, finds applications in industries ranging from pharmaceuticals for anti-inflammatory treatments to cosmetics for exfoliating and rejuvenating skincare products [10]. Hence, papaya has been selected as the focus of this project over other tropical fruits available in Malaysia.

The best hope for autonomously developing important and unique traits lies in deep learning approaches, particularly CNNs. A large number of convolutional layers are utilized in deep learning (DL) to capture features from datasets [11], [12]. Recognition of plant diseases may be achieved by a deep-learning model [13]. Deep learning has several disadvantages, since it needs substantial quantities of data for network training. The performance deteriorates when an available dataset lacks sufficient pictures. Transfer learning has set of benefits; for instance, it does not need a substantial volume of data to train the network. Transfer learning enhances the acquisition of a new task by using information obtained from a previously learnt, analogous activity. Numerous research used machine and deep learning in their methodologies for several fields such as humen medical [14], [15], geographical [16], [17], [18], agriculture [19], and etc. Transfer learning displays advantages such as less training duration, minimised generalisation error, and lower computing expenses in developing a deep learning model [20].

CNN models excel at handling image data and can automatically extract key features from images, such as color, texture, and shape, which are crucial for distinguishing the ripeness and freshness of papaya fruit [21]. Through training, the model can learn the subtle differences between good and bad fruits, achieving highly accurate classification. Using CNN models, unmarketable fruits can be quickly screened out in the early stages, ensuring that only high-quality papaya reaches the market, thereby enhancing consumer satisfaction. By using transfer learning, we may assign certain weights to models that have already been trained on massive picture datasets like VGG16, ResNet50, and InceptionV3. This enables us to apply these models to new tasks with much less data and computing resources required for training [22].

Therefore, the purpose of this paper is to construct multiple CNN and transfer learning models to classify papaya fruits as healthy or diseased (unsuitable for sale) by learning the implicit features in the image pixels, thereby reducing the need for manual sorting of papaya fruits.

The contributions of this work to the corpus of knowledge are as follows:

1. The study thoroughly evaluates four notable architectures: —Custom CNN, VGG16, ResNet50 and InceptionV3, on their efficacy in detecting papaya is healthy or diseased.
2. This study depends a The original photos were sourced from various sources, such as publications, journals, and internet sites as a dataset. Among the raw photographs gathered, there are a total of 181 fruit images and 200 leaf images.

2. RELATED WORKS

A. S. Chaithanya et al. presents a Resnet-50 architecture-based transfer learning model for damaged papaya leaf recognition and classification. The model makes use of a dataset consisting of 2159 pictures, which includes both healthy and diseased leaves, including examples of anthracnose, bacterial spot, curl, and ringspots. When tested on a 220 image dataset, the model outperforms several basic models, including CNN, VGG 16, Inception V3, ResNet-50, DenseNet 121, MobileNet V2, and EfficientNet B0. Its accuracy rate is an outstanding 87.95% [23].

V. Kumar et al. proposed approach using neural networks using convolution, CNNs, and Random Forest algorithms to classify leaf diseases of papaya fruit. The research assesses illness classification efficacy across various groups, including PRSV, Dusty Mildew, bacterial Leaves Spot, Cercospora The leaves Spot, and Angular Leaf Spot. The algorithm performs well, with precision values consistently exceeding 93% for each disease class. The high recall values of the model reach 92%, indicating its ability to detect true positive occurrences. The F1-score, which assesses the balance of recall and accuracy, frequently exceeds 94%, proving the persistence of the categorization model. The paper highlights the model's ability to differentiate between a variety of papaya leaf diseases with an accuracy of 94.49% [24].

T. A. Mir et al presented six sets of arranged plants according destect diseases in leaf papaya using Random Forest (RF) and Convolutional Neural Networks (CNN) "Healthy," "Papaya Ringspot Virus (PRSV)," "Papaya Leaf Curl Virus (PLCV)," "Powery Mildew," "Papaya Mosaic Virus (PapMV)," and "Leaf Spot Diseases." The concept classifies plants into six categories. The model evaluates its capacity to separate these categories using many criteria, including accuracy, recall, and F1-score. The accuracy of the model almost between 94.44% and 96.22%, which means it achieves proof positive result, and the recall is almost between 94.12% to 96.13%, which means it can recognize accurately between different situations, and F1-score 95.48%, that proof good perfomance and the last score is 95.49%, the model considers effictice in detect of papaya leaf diseases that be usefull in care of crops agricultural [25].

Suryavanshi and the team proposed an intelligent tool using a convolutional neural network and federated learning to find a solution to the spread problem of papaya leaf diseases. The system achieved an accuracy of 80.32% (for cy_1-fe_1) and 90.37% (for cy_5-fe_5) from local data for five types of papaya leaf diseases. They evaluated the model using different measures like accuracy, recall, and F1-score, which proved that it was a strong and flexible model across various versions. Cy_5 gave the best results [26].

S. Mehta et al. presented a new approach named Federated Learning (FL) using a Convolutional Neural Network (CNN) to detect disease in papaya leaves and classify it. Having five main papaya leaf diseases under their scope, the study split each into five degrees of disease severity. With accuracy ranging from 87% in powdery mildew to 92% in Papaya Leaf Distortion Mosaic Virus, the CNN model had an overall classification accuracy of 92%. The result implies that further work is required to enhance data management and model resilience in many agricultural environments [27].

The purpose of this study is to provide an AI-based solution to the problem of papaya plant disease in Malaysia's agricultural industry. We gathered 225 pictures, split them into three categories: Healthy, Ring Spot, and Black Spot, from various Google sources and real-life photos. We got 600 photos after doing all the necessary preprocessing and augmentation. After a 10-epoch run via Google Colab, InceptionV3 was determined to have the highest validation accuracy and was therefore selected as the best base model. In order to optimize the hyperparameters, we extracted 2048 output features from the final model layer and stored them in NumPy arrays. After 50 iterations using the SGD optimizer, a batch size of 4, and a learning rate of 0.01 the trained data attained a validation accuracy of 1.0. A validation accuracy of 0.96 and a training accuracy of 1.0 were produced by a fully trained model using training photos. In order to test the app's accuracy, we imported 15 photos from each class. The results showed an accuracy, recall, f1-Score, and precision of 0.8889 [28].

The study explores the use of Vision Transformers (ViTs) in deep learning for the precise identification of papaya diseases. The paper shows in terms of computational efficiency and classification accuracy ViTs' better performance than conventional convolutional neural networks (CNNs). Emphasizing the need of using sophisticated models like ViTs for disease control and sustainable agricultural practices, the research helps to progress computer vision methods in agriculture. With ViTs used in real-world situations, early disease identification, better crop management methods, and contribution to food security and sustainable agriculture acheaved 91% accuracy stands promise [29].

Using federated learning, CNN, and the FedAvg algorithm, S. Mehta and R. Kumar propose a novel method to classify four key illnesses impacting papaya plants and their leaves. Its accuracy is 91%; its recall is 90%; its F-Measure is 90%. With a 5% F1-score, the response accuracy overall to disease classification is 92%. By lowering presented data and privacy concerns, the federated learning approach decreases training bandwidth by up to 76% for various models [30].

Using federated learning and convolutional neural networks (CNN), S. Malhotra et al. offers a fresh approach for spotting and categorizing illnesses in papaya leaves. Using data from four customers, the study focuses on four degrees of bad health in papaya leaves. In identifying disorders, clients SV_1 and SV_2 shown varying performance. With a better weighted average and average score, Client SV_2 scored higher overall. Client SV_3 gave a good average and weighted score. With an average of 92.89 for all macro components and 92.74 for tiny micro parts linked to it, client SV_4 showed excellent performance [31].

Using a machine learning model, S. J. Banarase and S. D. Shirbahadurkar show a novel method for spotting illnesses in papaya fruits and leaves. The model uses hyperparatunning of machine learning classifiers along with grey level co-occurrence matrix (GLCM) feature extraction. Extensively using GLCM, the model analyzes eight classes of healthy and sick papaya leaf and fruit to extract sixteen attributes. More exact outcomes come from machine learning techniques such Random Forest (RF) and Hyperparatunning using Support Vector Machine (SVM). With SVM with hyperparatunning attaining 91.47% accuracy and RF at 90.22% , the method's general accuracy is higher than that of current approaches [32].

Table 1 summary studies:

Table 1: Summmery Related Works

Ref	Method	Measure	Result	Finding
[23]	Transfer Learning (ResNet-50)	Accuracy	87.95%	Outperformed base models (CNN, VGG16, etc.) in classifying 5 diseases.
[24]	CNN + Random Forest	Precision, Recall, F1-score	Precision >93%, Recall 92%, F1 >94%	High efficacy in classifying PRSV, mildew, and leaf spots.
[25]	CNN + Random Forest	Accuracy, Recall, F1-score	Accuracy 94.44–96.22%, F1 95.48%	Reliable for 6 disease classes; suitable for real-world farming.
[26]	Federated Learning + CNN	Precision, Accuracy	Precision 80–90% (varies by client)	Decentralized approach with robust metrics across clients.
[27]	Federated Learning + CNN	Accuracy	92% (range: 87–92% per disease)	FL effective but needs robustness improvements.
[28]	InceptionV3 (Transfer Learning)	Precision, Recall, F1, Accuracy	All metrics 88.89%, Val. accuracy 96%	High accuracy but potential overfitting (training accuracy 100%).
[29]	Vision Transformers (ViTs)	Accuracy	91%	ViTs outperformed CNNs in accuracy and computational efficiency.
[30]	FL + CNN + FedAvg	Accuracy, Recall, F1	Accuracy 91%, Recall 90%, F1 90%	Privacy-preserving; reduced training bandwidth by 76%.
[31]	Federated Learning + CNN	Weighted Avg. Scores	Client SV_4: 92.89% (macro), 92.74% (micro)	Client-specific performance variability in disease classification.
[32]	GLCM + SVM/RF (Hyperparameter Tuned)	Accuracy	SVM: 91.47%, RF: 90.22%	GLCM feature extraction + tuned classifiers outperformed existing methods.

3. METHODOLOGY

In this section, we first address data collection, data preprocessing, and the design of our own CNN network model, as well as introduce the architecture of three transfer models that we have adopted.

3.1 Dataset Collection

To diagnose Papaya Ringspot Virus (PRSV) disease, a well-prepared dataset containing pictures of the papaya tree's affected areas is necessary. This research focuses on the papaya fruit and leaves. The preparation procedures used to obtain the dataset for use in Convolutional Neural Network (CNN) model and transfer learning techniques will be discussed here. The two major sections of the papaya tree, the leaf and the fruit, impacted by PRSV have been assigned as separate labels. The original photos were sourced from various sources, such as publications, journals, and internet sites. Among the raw photographs gathered, there are a total of 181 fruit images and 200 leaf images. Figure 2 shows the samples of dataset original images of papaya fruit with ringspot virus.

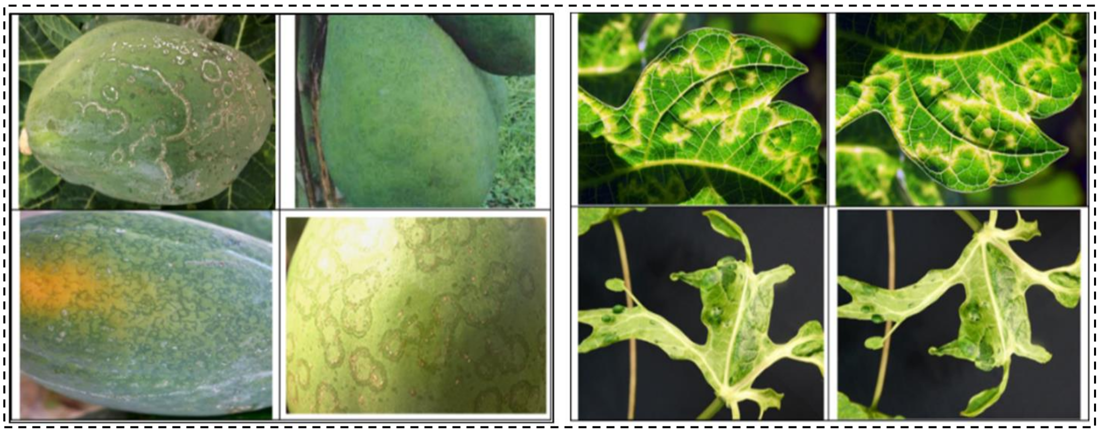

Figure 2: Sample of Dataset

3.2 Preprocessing

To enhance the dataset and the model's reliability, data augmentation techniques have been employed. Three different kinds of transformations were applied to each image: rotation, horizontal flipping, and vertical flipping. The dataset's size was practically quadrupled during this augmentation process. Following augmentation, there are a total of 724 photos of fruits and 800 photos of leaves.

For this method, an online application was utilized to rotate the images in bulk. The rotations included turning the images to the right, down, and left from the raw image orientation. This method tripled the quantity of the images acquired earlier, resulting in a total of 1,524 photos being prepared for the research as show in Figure 3 the diffrent of number before and after augumentation way.

Once the raw images for both the fruit and leaf labels were finalized, the images were resized to a resolution of 224x224 pixels. For this resizing process, an online application capable of bulk resizing was utilized to efficiently resize 1,500 raw images.

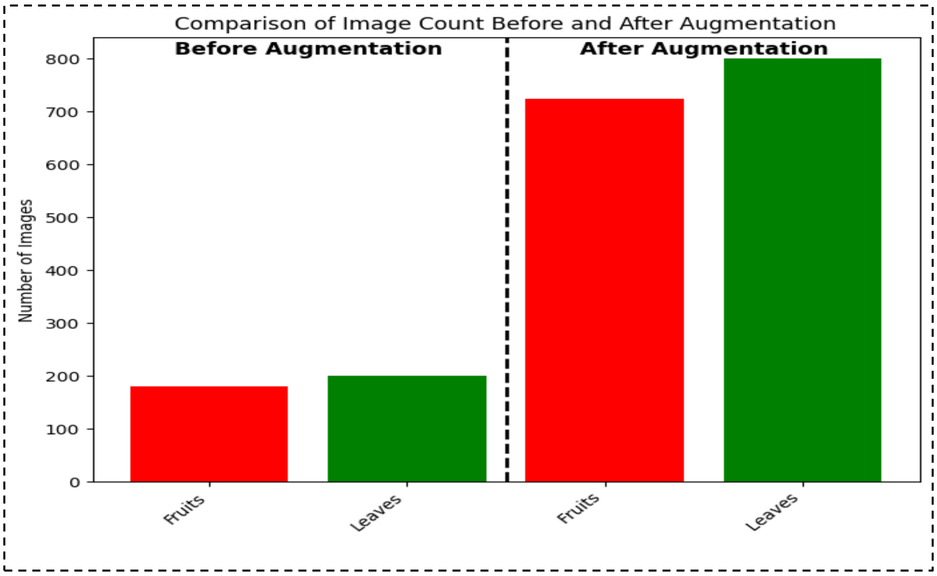

Figure 3: The diffrent number of Images before and after augumentation

3.3 CNN Custom

In this study, a custom Convolutional Neural Network (CNN) model is developed to classify papaya ringspot virus disease into two categories: fruit and leaf. The architecture of the proposed CNN model is illustrated in Figure 4.

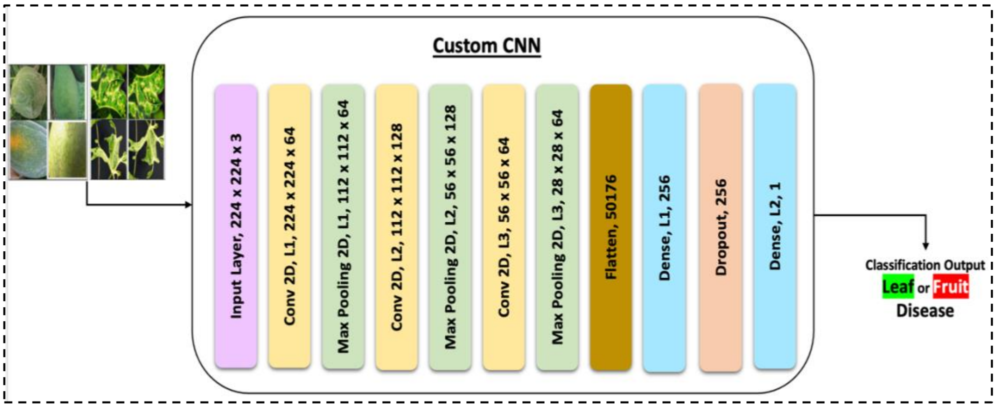

Figure 4: Architecture of CNN Custom

Ten layers make up the convolutional neural network (CNN) model: output, max pooling, flattening, dense, and dropout. To improve computational efficiency, the input pictures are reduced in spatial dimensions in the first six layers, which consist of alternating convolutional and max pooling layers. These layers gradually extract hierarchical features. The CNN is able to distinguish between unhealthy leaves and fruits because it employs filters of varying sizes to detect relevant patterns in the dataset.

After the convolutional layers, the flatten layer convert the found feature maps into a one line array, making them suitable for fully connected layers. The first dense layer, containing 256 neurons, processes these features to learn complex representations, followed by a dropout layer that mitigates overfitting by randomly deactivating some neurons during training. Finally, the output layer, consisting of a single neuron with a sigmoid activation function, determines whether the given image represents a diseased fruit or leaf.

This CNN structure is created to enhance classification execution while keep computational efficiency, ensuring robust and accurate classification of papaya ringspot virus diseases.

3.4 Transfer Learning

3.4.1 VGG16

In order to classify papaya fruits and leaves as containing the Papaya Ringspot Virus (PRSV), this paper makes use of transfer learning using the pre-trained VGG16 model. The Visual Geometry Group at the University of Oxford developed VGG16, which is a strong contender for illness diagnosis due to its exceptional performance in picture classification challenges [33].

The main aim of using VGG16 is to optimize both the efficiency and accuracy of PRSV classification by capitalizing on its deep hierarchical feature extraction capability. This model is well-known for its structured simplicity and substantial depth. It is designed to process images of size 224×224 pixels in the RGB color space. During training, the mean RGB value of each image is computed and fed into the VGG16 convolutional network to ensure consistent preprocessing.

In order to capture complex visual details, VGG16 has 13 convolutional layers that use 3×3 kernel filters. The total number of layers in VGG16 is 16. In addition, max pooling layers are used to minimize computational complexity and preserve vital information while reducing the spatial dimensions of feature maps. As a last step in the classification process, fully linked layers distribute the collected characteristics among the target classes as a probability. You may see VGG16's architecture in Figure 5.

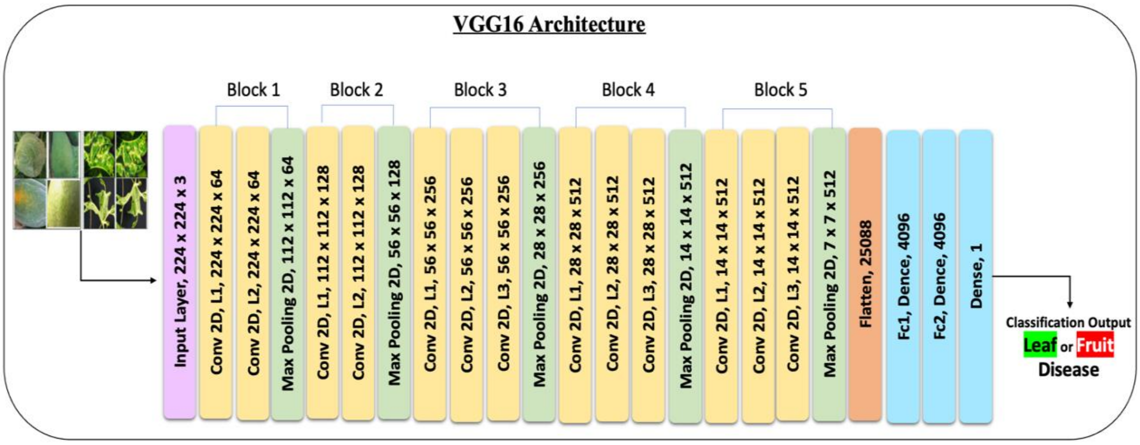

Figure 5: Architecture of VGG16

3.4.2 ResNet50

Another proposed approach utilizes the ResNet50 pre-trained model to harness the power of transfer learning for identifying PRSV disease. ResNet50, developed by Kaiming He et al. at Microsoft Research Asia in 2015, is a deep convolutional neural network that has revolutionized deep learning. It is acclaimed for its capability to train exceptionally deep networks with hundreds of layers using residual blocks and skip connections. These architectural features enable ResNet50 to retain information from earlier layers, facilitating the learning of more effective representations from input data [34].

By utilizing ResNet50, the goal is to improve the efficiency and accuracy of PRSV detection in papaya fruits and leaves. The model takes a 224x224 RGB image as input, which undergoes preprocessing involving the normalization of its RGB values before entering the ResNet50 convolutional network. ResNet50 comprises 50 layers organized into 5 stages, each containing a series of residual blocks used in this study. The architecture addresses the challenge of training deep networks by incorporating skip connections, also known as identity connections. These connections facilitate more effective gradient flow through the network, thereby enabling the training of deeper models.

network begins with Stage 1 (Layer 0 to Layer 6), initiating with initial convolutions, batch normalization, ReLU activation, and max-pooling to extract fundamental features from the input image. Stage 2 (conv2_x, Layer 7 to Layer 38) incorporates three residual blocks that extract mid-level features, utilizing pre-trained weights for enhanced generalization. Stage 3 (conv3_x, Layer 39 to Layer 80) includes four residual blocks for intermediate feature extraction, capturing more intricate patterns. Stage 4 (conv4_x, Layer 81 to Layer 142), featuring six residual blocks, focuses on advanced feature extraction, capturing detailed characteristics relevant to PRSV. Stage 5 (conv5_x, Layer 143 to Layer 174) encompasses three residual blocks that refine data for classification, fine-tuning high-level features. The final layers involve average pooling, flattening, and passing through a fully connected layer and softmax layer to generate the final probability distribution across classes, determining whether the papaya leaf is infected with PRSV.

3.4.3 InceptionV3

The proposed work utilizes InceptionV3 pre-trained model to leverage transfer learning advantages in detecting Papaya Ringspot Virus (PRSV). InceptionV3 is an advanced and efficient deep learning model designed for image classification and object detection [35] . By harnessing its cutting-edge architectural features, this study aims to achieve high accuracy in PRSV detection on papaya fruits and leaves while maintaining computational efficiency.

InceptionV3, developed by Google, represents a significant advancement in convolutional neural network architectures, evolving from earlier models like Inception (GoogLeNet) and InceptionV2 [36]. It incorporates features such as Inception modules, which utilize parallel convolutions of varying sizes to capture multi-scale features, and factorized convolutions to reduce computational costs without sacrificing accuracy. The model also integrates extensive batch normalization to stabilize and accelerate training, auxiliary classifiers to enhance gradient flow and convergence, efficient grid size reduction techniques, and asymmetric convolutions to further minimize computational load.

The architecture of InceptionV3 includes initial layers for preprocessing, stacked Inception modules for different abstraction levels, reduction modules to manage complexity, auxiliary classifiers for regularization, and final layers for classification [37]. Known for its high accuracy, computational efficiency, and adaptability for transfer learning, InceptionV3 proves suitable for diverse applications ranging from medical imaging to autonomous driving.

4. Experiment

Based on the proposed methodology outlined in the previous chapter, this section presents the experimental setup, performance results, and comprehensive analysis of the models used for Papaya Ringspot Virus (PRSV) detection. The experiments were conducted using different configurations for each model to evaluate their effectiveness in classifying infected papaya fruit and leaf images. The findings are discussed in detail to highlight learning behavior, generalization performance, and comparative strengths across the tested models.

4.1 Experiment Setup

The experiment is conducted on images of papaya fruits and leaves affected by PRSV disease. Firstly, the dataset is divided into training 70%, validation 20% and testing 10% sets. Training and validation are performed concurrently. The Python library "Keras" is utilized to implement the CNN model and the three pre-trained models on Google Colab.

The experimental setup involves training and evaluating various convolutional neural network (CNN) architectures for PRSV detection. The models tested include CNN, VGG16, ResNet50, and InceptionV3. All models utilize the ReLU activation function, with CNN and VGG16 also incorporating Sigmoid for final layers. Adam optimizer with a learning rate of 0.001 is consistently used across all experiments. Each model undergoes initial training for 10 epochs, with some configurations involving the unfreezing of specific layers for fine-tuning purposes to enhance model performance.

Figure 6 explains the models used along with the number of parameters they generated. This shows that VGG16 is capable of generating more parameters compared to CNN and the other pre-trained models.

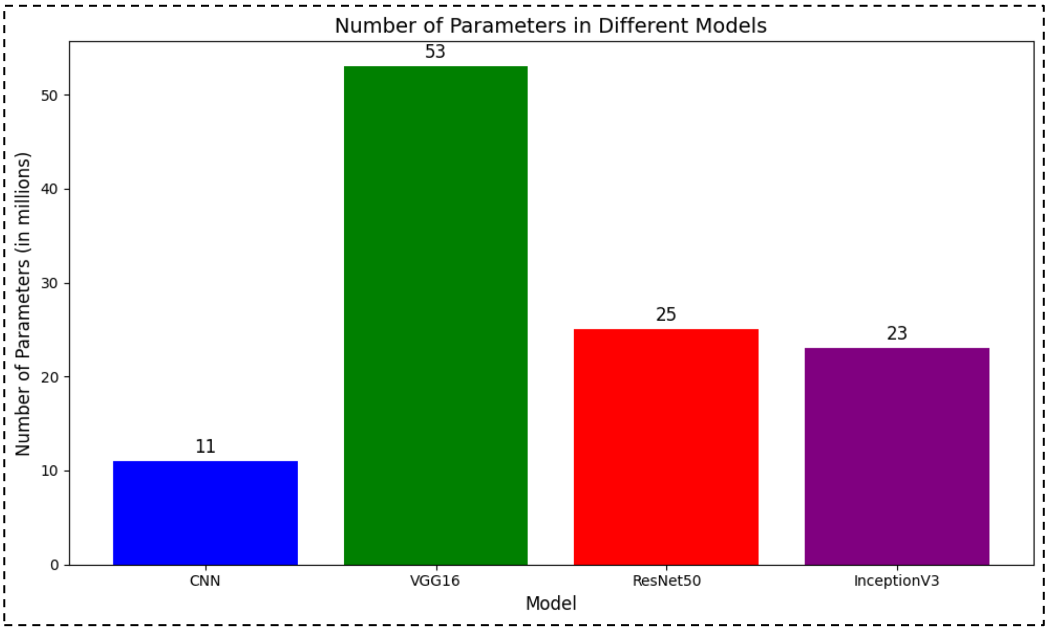

Figure 6: Model and Number of Parameters Generated

The CNN, with 11 million parameters, is trained for 10 epochs without any fine-tuning. VGG16 and ResNet50 are initially trained for 10 epochs, after which the last four layers are unfrozen for further training, aiming to refine the models' learning by allowing them to adjust weights in the final layers. ResNet50 is also evaluated with an extended training period of 25 epochs without fine-tuning to observe the impact of prolonged training on model accuracy. Similarly, InceptionV3 is tested in multiple configurations: a 10-epoch training with and without fine-tuning, and a 25-epoch training to assess the effects of extended training and fine-tuning on its performance.

Model performance is measured by accuracy at the first epoch and final accuracy after the complete training period. Initial accuracy indicates the models' early learning capabilities, while final accuracy reflects their generalization ability after full training. The comparison also considers the number of parameters in each model, which varies significantly and influences computational efficiency and performance.

4.2 Experimental Results

4.2.1 Result of Custom CNN

The Adam optimizer was used to train the custom Convolutional Neural Network (CNN) model for 10 epochs at a learning rate of 0.001. For binary classification, the design made use of a Sigmoid function in the output layer and the ReLU activation function in the hidden layers. The model was trained and built from the ground up without the use of any transfer learning or fine-tuning. The CNN model has the fewest trainable parameters of all the models tested here, with only under 11 million total.

At epoch 1, the model achieved a training accuracy of 81.91%, which steadily increased to 99.22% by epoch 10. The corresponding validation accuracy started at 75.2% and improved to 89.78% over the same training period. This gradual improvement indicates that the model successfully learned relevant features, although its generalization to unseen data remained limited when compared to the transfer learning models.

Figure 7 shows that both the training loss and the validation loss fell steadily from 0.88 to 0.11 and 0.97 to 0.28, respectively. Overfitting, which occurs often in models trained from beginning without regularization or pre-learned information, may be at play here, as shown by the discernible performance disparity between training and validation measures.

Despite these limitations, the custom CNN model demonstrated effective learning behavior and served as a valuable baseline. Its performance validates the potential of a lightweight, scratch-trained model for PRSV detection, while also highlighting the advantages of leveraging transfer learning for improved generalization.

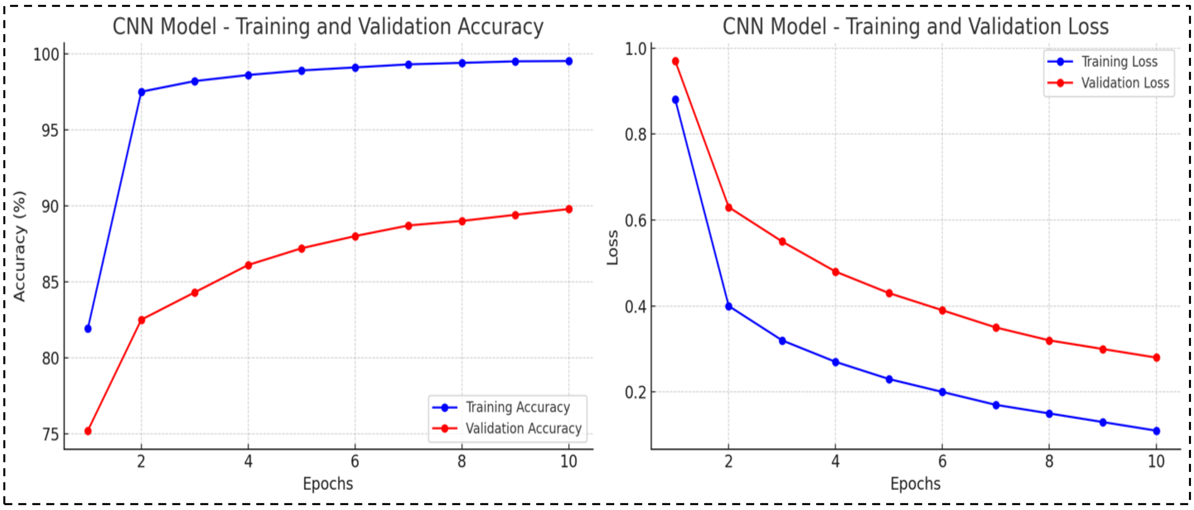

Figure 7: Curve of Training and Validation for Accuracy and Loss (CNN Custom)

4.2.2 Result of VGG16

The VGG16 model was trained for 10 epochs using the Adam optimizer with a learning rate of 0.001. The architecture incorporated the ReLU activation function in its intermediate layers and a Sigmoid activation in the output layer to support binary classification. As part of the training process, the final four layers of the network were unfrozen for fine-tuning, allowing the model to better adapt pre-trained ImageNet features to the specific task of PRSV detection in papaya fruit and leaf images.

At the start of training, VGG16 achieved a training accuracy of 93.71%, which increased rapidly to 100% by epoch 10. The validation accuracy also showed consistent improvement, starting at 91.5% and reaching 100% in the final epoch. This close alignment between training and validation accuracy indicates excellent generalization performance, with minimal risk of overfitting.

Training loss fell from 0.41 to 0.01, while validation loss fell from 0.45 to 0.01; both curves decreased smoothly, as shown in Figure 8. The training and validation curves closely coincide across the epochs, which shows that the model converges well and learns robust features from the data.

VGG16 is the largest model evaluated in this study, comprising approximately 53 million trainable parameters. Despite its computational demands, the model demonstrated strong performance and stability. The combination of deep architecture, pre-trained weights, and targeted fine-tuning contributed to its outstanding accuracy and convergence, confirming its suitability for high-precision PRSV detection tasks.

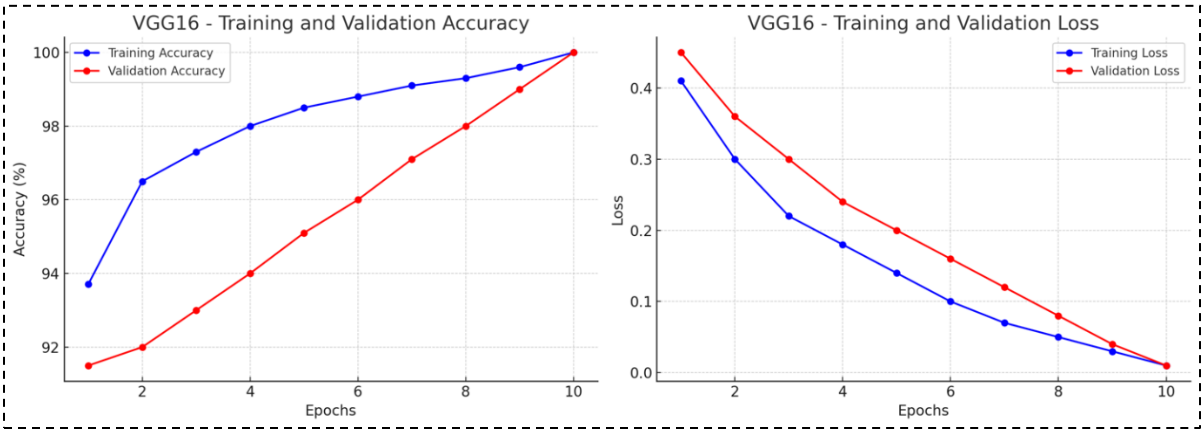

Figure 8: Curve of Training and Validation for Accuracy and Loss (VGG16)

4.2.3 Result of ResNet50

The ResNet50 model was evaluated under three different configurations to assess its performance in PRSV detection:

(i) Training for 10 epochs without fine-tuning,

(ii) Training for 10 epochs with fine-tuning of the last four layers, and

(iii) Extended training for 25 epochs without fine-tuning.

In all configurations, the model used the ReLU activation function and the Adam optimizer with a learning rate of 0.001. The ResNet50 architecture comprises approximately 25 million parameters.

The first setup included training the model for 10 epochs without thawing any layers. Its validation accuracy was 84.0 percent, while its initial training accuracy was 87.13%. Both measures were 100% at the tenth epoch. The validation loss fell from 0.49 to 0.01 and the training loss from 0.43 to 0.01. Even though this configuration made use of pre-trained weights, the slower adaptation to the PRSV dataset is shown by the substantially lower beginning validation accuracy and larger initial loss.

In the second configuration, the last four layers of ResNet50 were unfrozen for fine-tuning. This resulted in an improved initial training accuracy of 92.12% and validation accuracy of 86.2%, both of which steadily increased to 100% by epoch 10. Training and validation losses also dropped rapidly, from 0.39 to 0.01 and 0.45 to 0.01, respectively. This highlights the impact of fine-tuning in enabling the model to adjust its higher-level features for better task-specific learning and generalization.

The third configuration involved extended training for 25 epochs without fine-tuning. The model began with an initial training accuracy of 92.36% and validation accuracy of 84.0%, and both reached 100% by the end. The training loss declined from 0.42 to 0.01, and the validation loss from 0.48 to 0.01 over 25 epochs. While the model eventually achieved excellent performance, it required more training time compared to the fine-tuned configuration.

As illustrated in Figure 9, all configurations of ResNet50 converged effectively. However, the 10-epoch fine-tuned setup offered the best balance between speed, accuracy, and loss reduction, making it the most practical and efficient configuration for PRSV detection in papaya leaves and fruits.

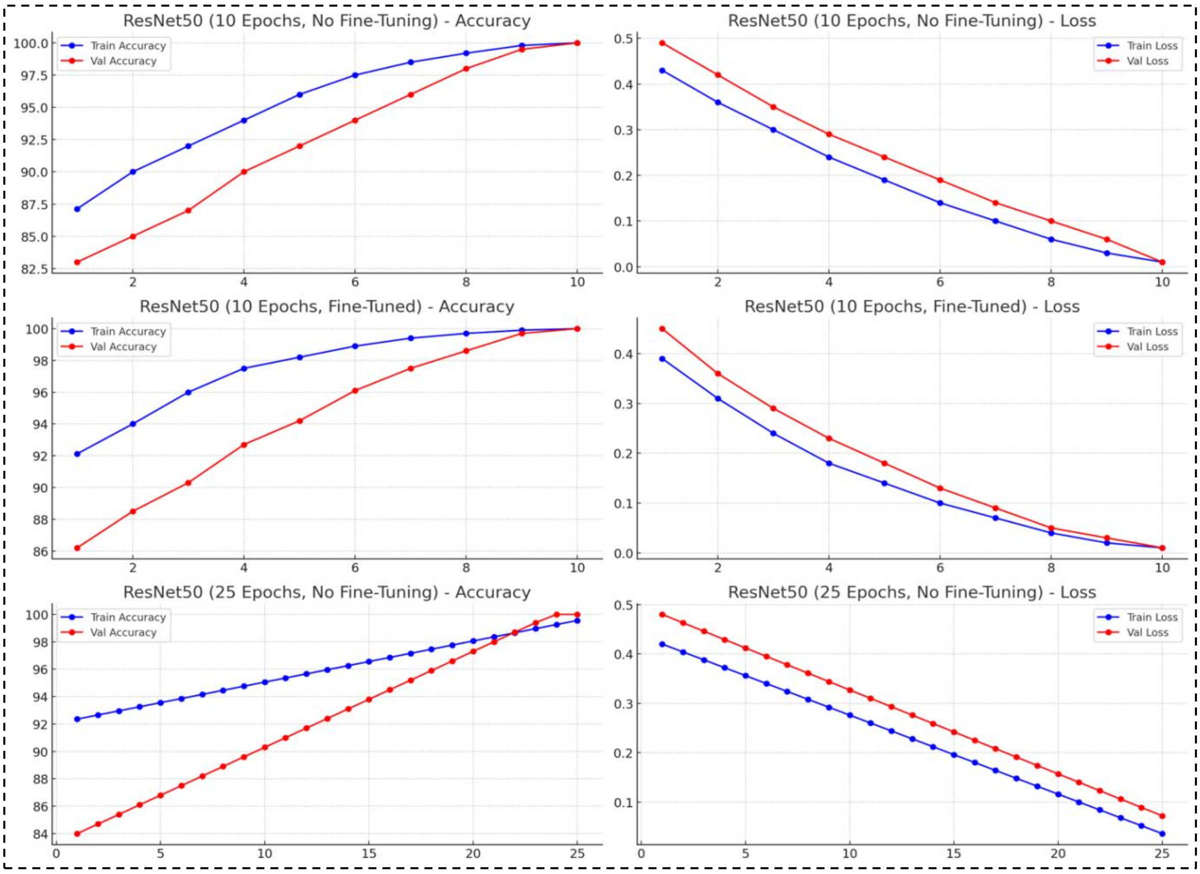

Figure 9: Curve of Training and Validation for Accuracy and Loss (ResNet50)

4.2.4 Result of InceptionV3

The InceptionV3 model was evaluated using three configurations to investigate its performance in detecting PRSV from papaya fruit and leaf images:

(i) Training for 10 epochs without fine-tuning,

(ii) Training for 10 epochs with fine-tuning applied to the last four layers, and

(iii) Extended training for 25 epochs without fine-tuning.

All configurations utilized the ReLU activation function and were optimized using the Adam algorithm with a learning rate of 0.001. InceptionV3 is the most lightweight among the tested pre-trained models, with approximately 23 million trainable parameters.

The first setup included training the model for 10 epochs without making any adjustments to its parameters. Starting with a validation accuracy of 90% and a training accuracy of 92.52%, it eventually reached 100% in both measures by the last epoch. There was a decrease from 0.31 to 0.01 in training loss and from 0.38 to 0.01 in validation loss. This setup proved that the model could generalize well even without tweaking.

In the alternate setup, the last four layers were adjusted to perfection. As a result, by the tenth epoch, the validation accuracy had increased to 94.5 percent and the initial training accuracy to 96.02%. Training losses decreased from 0.22 to 0.005, while validation losses decreased from 0.28 to 0.005. The efficiency of targeted layer updates was further reinforced by the fact that this fine-tuned configuration converged faster and showed somewhat greater generalization than the non-fine-tuned version as shown Figure 10.

The model underwent 25 epochs of training in the third configuration without any fine-tuning. Its validation accuracy was 93.1 percent and its training accuracy was 95.3 percent; by epoch 25, both had improved to 100 percent. The validation loss dropped from 0.34 to 0.01 and the training loss from 0.29 to 0.01 throughout the prolonged training period. The training time was longer and more computationally intensive, but the outcomes were identical to the fine-tuned model.

As illustrated in Figure 10, all three InceptionV3 configurations performed exceptionally well, achieving perfect accuracy with minimal loss by the end of training. Among them, the 10-epoch fine-tuned configuration offered the best trade-off between training time, early performance, and generalization, making it the most optimal setup for efficient and accurate PRSV detection.

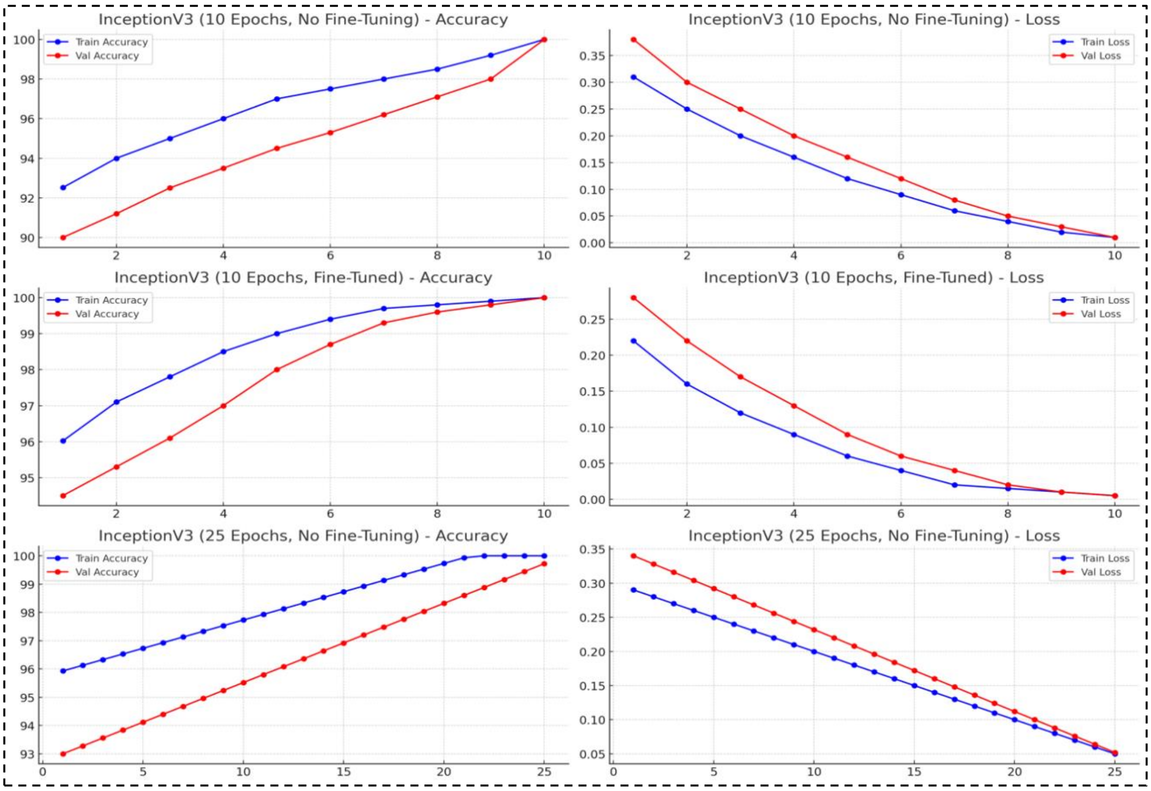

Figure 10: Curve of Training and Validation for Accuracy and Loss (InceptionV3)

4.3 Experimental Discussion

When comparing the CNN to the pre-trained models, the custom CNN, with 11 million parameters, starts with an initial accuracy of 81.91% and reaches 99.22% after 10 epochs. While this performance is commendable, it falls short of the perfect accuracy achieved by the pre-trained models. This discrepancy highlights the advantage of pre-trained models, which have been trained on extensive datasets, enabling them to generalize better to new tasks. The CNN, though effective, lacks the comprehensive feature knowledge that pre-trained models like InceptionV3, VGG16, and ResNet50 possess through transfer learning.

The pre-trained models appear to have good performance and efficiency, especially the inceptionV3 accuracy reaches 100% with parameters less than VGG16, that proof enhancing in design. VGG16 model also achieves good performance, though it has many parameters that mean more mathematical complexity. ResNet50 appears to continue good performance that reaches accuracy to perfect without parameter tuning and long training, all of which explains the strong performance of transfer learning models. Generally, the results indicate that transfer learning models are accurate, effective, and suitable for detecting PRSV in papaya fruits and leaves. Table 2 explains the overall results of the models in this study.

Additionally, the comparison of loss for training and validation support via models shows these results. The custom CNN registers a high loss in validation and notices a gap in training loss. Other hand, the pre-trained models appear same values for each training and validation, with faster enhancing, which indicates that pre-trained models generalize better and are less overfitting. The results also ensure that the pre-trained models are trust approach to create effective and accurate models for PRSV detection.

Table 2: Model Performance Summary

Model	Parameters (M)	Init. Train Acc (%)	Final Train Acc (%)	Init. Val Acc (%)	Final Val Acc (%)	Init. Train Loss	Final Train Loss	Init. Val Loss	Final Val Loss
CNN	11	81.91	99.22	75.2	89.78	0.88	0.11	0.97	0.28
VGG16 (10 Epochs, Fine-Tuned)	53	93.71	100.0	91.5	100.0	0.41	0.01	0.45	0.01
ResNet50 (10 Epochs, No FT)	25	87.13	100.0	83.0	100.0	0.43	0.01	0.49	0.01
ResNet50 (10 Epochs, Fine-Tuned)	25	92.12	100.0	86.2	100.0	0.39	0.01	0.45	0.01
ResNet50 (25 Epochs, No FT)	25	92.36	100.0	84.0	100.0	0.42	0.01	0.48	0.01
InceptionV3 (10 Epochs, No FT)	23	92.52	100.0	90.0	100.0	0.31	0.01	0.38	0.01
InceptionV3 (10 Epochs, Fine-Tuned)	23	96.02	100.0	94.5	100.0	0.22	0.005	0.28	0.005
InceptionV3 (25 Epochs, No FT)	23	95.93	100.0	93.0	100.0	0.29	0.01	0.34	0.01

5. CONCLUSION AND FUTURE WORKS

To conclude, the detection of PRSV on papaya fruits and leaves using CNN and pre-trained models (VGG16, ResNet50, InceptionV3) can achieve successful predictions demonstrating high final accuracy across the experiments. However, the pre-trained models consistently outperformed the CNN model in detecting PRSV on papaya fruits and leaves, with all pre-trained models achieving a final accuracy of 100%. Additionally, the initial accuracy of the pre-trained models was superior to that of the CNN model.

The successful detection of PRSV using the pre-trained models can significantly benefit the agricultural industry in Malaysia by enabling early detection of real-life scenarios of PRSV disease in papaya plants, thereby potentially reducing losses in papaya production. The proposed models could facilitate more frequent and effective early detection efforts.

For future work, expanding the dataset with additional images could enhance the models' ability to recognize various types of images that might go undetected during testing. Ensuring that the models can accurately interpret all potential papaya images containing PRSV disease will be crucial for enhancing reliability and trustworthiness. Additionally, integrating these models into mobile or web-based applications could provide real-time diagnostic tools for farmers and agricultural officers. Field-testing and user validation will be important future steps to ensure the models' usability and reliability under practical conditions. These efforts will pave the way toward intelligent, data-driven disease management solutions in agriculture.

REFERENCES

[1] Plant pests and diseases. [EB/OL]. [2022-03-09]. http://www.fao.org/ emergencies/emergency-types/plant-pests-and-diseases/en/.

[2] J. Liu and X. Wang, “Plant diseases and pests detection based on deep learning: a review,” Plant Methods, vol. 17, no. 1, p. 22, Dec. 2021, doi: 10.1186/s13007-021-00722-9.

[3] A. Singh et al., “Challenges and Opportunities in Machine-Augmented Plant Stress Phenotyping,” Trends in Plant Science, vol. 26, no. 1, pp. 53–69, Jan. 2021, doi: 10.1016/j.tplants.2020.07.010.

[4] R. Sekeli, M. H. Hamid, R. A. Razak, C.-Y. Wee, and J. Ong-Abdullah, “Malaysian Carica papaya L. var. Eksotika: Current Research Strategies Fronting Challenges,” Frontiers in Plant Science, vol. 9, Sep. 2018, doi: 10.3389/fpls.2018.01380.

[5] D. C. Mederos et al., “Ecology and evolution of papaya ringspot virus,” in Plant RNA Viruses, Elsevier, 2023, pp. 229–243. doi: 10.1016/B978-0-323-95339-9.00014-4.

[6] Y. K. Chan, “BREEDING PAPAYA FOR RESISTANCE TO RINGSPOT VIRUS DISEASE IN MALAYSIA,” Acta Horticulturae, no. 632, pp. 57–64, Feb. 2004, doi: 10.17660/ActaHortic.2004.632.6.

[7] S. R. Department, "Agriculture in Malaysia - statistics & facts," Statista Research Department, 7 Jun 2024. [Online]. Available: https://www.statista.com/topics/10680/agriculture-in-malaysia/. [Accessed 20 Jun 2024].

[8] H. N. Ngugi, A. E. Ezugwu, A. A. Akinyelu, and L. Abualigah, “Revolutionizing crop disease detection with computational deep learning: a comprehensive review,” Environmental Monitoring and Assessment, vol. 196, no. 3, p. 302, Mar. 2024, doi: 10.1007/s10661-024-12454-z.

[9] C. Chuwa and A. K. Dhiman, “Ripe Papaya: Nutrition and Health Benefits,” in Emerging Challenges in Agriculture and Food Science Vol. 6, Book Publisher International (a part of SCIENCEDOMAIN International), 2022, pp. 56–64. doi: 10.9734/bpi/ecafs/v6/3235B.

[10] D. o. Biosafety, Biology of Papaya (Carica papaya L.), Putrajaya: Ministry of Water, Land and Natural Resources, 2019.

[11] S. H. Lee, C. S. Chan, S. J. Mayo, and P. Remagnino, “How deep learning extracts and learns leaf features for plant classification,” Pattern Recognition, vol. 71, pp. 1–13, Nov. 2017, doi: 10.1016/j.patcog.2017.05.015.

[12] M. P. Shah, S. Singha, and S. P. Awate, “Leaf classification using marginalized shape context and shape+texture dual-path deep convolutional neural network,” in 2017 IEEE International Conference on Image Processing (ICIP), IEEE, Sep. 2017, pp. 860–864. doi: 10.1109/ICIP.2017.8296403.

[13] Y. Li and J. Yang, “Few-shot cotton pest recognition and terminal realization,” Computers and Electronics in Agriculture, vol. 169, p. 105240, Feb. 2020, doi: 10.1016/j.compag.2020.105240.

[14] A. Abed Mohammed, P. Sumari, and kassem Attabi, “Hybrid K-means and Principal Component Analysis (PCA) for Diabetes Prediction,” International Journal of Computing and Digital Systems, vol. 15, no. 1, pp. 1719–1728, Jun. 2024, doi: 10.12785/ijcds/1501121.

[15] L. S. Ashour et al., “Non-overlapping Patch-Based Pre-trained CNN for Breast Cancer Classification,” Iraqi Journal for Computer Science and Mathematics, vol. 6, no. 2, Jun. 2025, doi: 10.52866/2788-7421.1271.

[16] M. M. Abd Zaid, A. A. Mohammed, and P. Sumari, “Classification of Geographical Land Structure Using Convolution Neural Network and Transfer Learning,” Journal of Computer Science, vol. 20, no. 12, pp. 1580–1592, Dec. 2024, doi: 10.3844/jcssp.2024.1580.1592.

[17] M. M. Abd Zaid, A. A. Mohammed, and P. Sumari, “Remote Sensing Image Classification Using Convolutional Neural Network (CNN) and Transfer Learning Techniques,” Journal of Computer Science, vol. 21, no. 3, pp. 635–645, Mar. 2025, doi: 10.3844/jcssp.2025.635.645.

[18] M. M. Abd Zaid, A. Abed Mohammed, and P. Sumari, “Classification of Road Features Using Convolutional Neural Network (CNN) and Transfer Learning,” International Journal of Computing and Digital Systems, vol. 17, no. 1, pp. 1–12, Jan. 2025, doi: 10.12785/ijcds/1571031764.

[19] B. H. Jawad, O. A. Alwesabi, N. Abdullah, and A. A. Mohammed, “Hybrid Artificial Neural Network Activation Function to Reduce Water Wastage in Agricultural Irrigation,” IEEE Access, vol. 13, pp. 93302–93322, 2025, doi: 10.1109/ACCESS.2025.3573678.

[20] M. Hussain, J. J. Bird, and D. R. Faria, “A Study on CNN Transfer Learning for Image Classification,” 2019, pp. 191–202. doi: 10.1007/978-3-319-97982-3_16.

[21] G. Shrestha, Deepsikha, M. Das, and N. Dey, “Plant Disease Detection Using CNN,” in 2020 IEEE Applied Signal Processing Conference (ASPCON), IEEE, Oct. 2020, pp. 109–113. doi: 10.1109/ASPCON49795.2020.9276722.

[22] Ch. Mukesh, A. Likhita, and A. Yamini, “Performance Analysis of InceptionV3, VGG16, and Resnet50 Models for Crevices Recognition on Surfaces,” 2024, pp. 161–172. doi: 10.1007/978-981-99-7817-5_13.

[23] A. S. Chaithanya and M. Rachana, “Identification of Diseased Papaya Leaf through Transfer Learning,” Indian Journal Of Science And Technology, vol. 16, no. 48, pp. 4676–4687, Dec. 2023, doi: 10.17485/IJST/v16i48.2690.

[24] V. Kumar, D. Banerjee, D. Upadhyay, M. Singh, and K. R. Chythanya, “Hybrid Model for EffectivePapaya Leaf Disease Diagnosis,” in 2024 International Conference on E-mobility, Power Control and Smart Systems (ICEMPS), IEEE, Apr. 2024, pp. 01–05. doi: 10.1109/ICEMPS60684.2024.10559298.

[25] T. A. Mir, D. Banerjee, M. Kumar, R. Rawat, and Y. Chanti, “Hybridized Model for Improved Papaya Leaf Disease Classification: CNN and Random Forest Integration,” in 2024 5th International Conference for Emerging Technology (INCET), IEEE, May 2024, pp. 1–6. doi: 10.1109/INCET61516.2024.10593105.

[26] A. Suryavanshi, V. Kukreja, S. Chamoli, S. Mehta, and A. Garg, “A New Age for Agricultural Diagnostics: Federated CNN for Papaya Leaf Diseases,” in 2024 IEEE International Conference on Interdisciplinary Approaches in Technology and Management for Social Innovation (IATMSI), IEEE, Mar. 2024, pp. 1–6. doi: 10.1109/IATMSI60426.2024.10502474.

[27] S. Mehta and P. Sharma, “Implementing CNNs with Federated Learning for Enhanced Detection of Papaya Leaf Diseases,” in 2024 4th Asian Conference on Innovation in Technology (ASIANCON), IEEE, Aug. 2024, pp. 1–5. doi: 10.1109/ASIANCON62057.2024.10837810.

[28] B. X. Wong, C. C. Chew, K. Gaik Tay, O. Al-Qershi, A. Huong, and S. A. Alzaeemi, “Android Based-App Papaya Leaf Disease Identification Using Convolution Neural Network,” 2023. [Online]. Available: http://arqiipubl.com/ams

[29] Monali Parmar and Dr. Sheshang Degadwala, “Deep Learning for Accurate Papaya Disease Identification Using Vision Transformers,” International Journal of Scientific Research in Computer Science, Engineering and Information Technology, vol. 10, no. 2, pp. 420–426, Mar. 2024, doi: 10.32628/CSEIT2410235.

[30] S. Mehta and R. Kumar, “Innovative Approaches in Agri-tech: Federated Learning Models for Disease Detection in Papaya Leaves,” in 2024 4th Asian Conference on Innovation in Technology (ASIANCON), IEEE, Aug. 2024, pp. 1–5. doi: 10.1109/ASIANCON62057.2024.10837859.

[31] S. Malhotra, A. Garg, V. Kukreja, and S. Mehta, “Agricultural AI: Integrating Federated Learning and CNN in Papaya Leaf Disease Detection,” in 2024 3rd International Conference for Innovation in Technology (INOCON), IEEE, Mar. 2024, pp. 1–6. doi: 10.1109/INOCON60754.2024.10512288.

[32] S. J. Banarase and S. D. Shirbahadurkar, “Papaya Diseases Detection Using GLCM Feature Extraction and Hyperparatuning of Machine Learning Approach,” 2023, pp. 145–158. doi: 10.1007/978-981-19-7874-6_12.

[33] Z.-P. Jiang, Y.-Y. Liu, Z.-E. Shao, and K.-W. Huang, “An Improved VGG16 Model for Pneumonia Image Classification,” Applied Sciences, vol. 11, no. 23, p. 11185, Nov. 2021, doi: 10.3390/app112311185.

[34] K. He, X. Zhang, S. Ren, and J. Sun, “Deep Residual Learning for Image Recognition,” in 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, Jun. 2016, pp. 770–778. doi: 10.1109/CVPR.2016.90.

[35] M. Z. Alom, M. Hasan, C. Yakopcic, T. M. Taha, and V. K. Asari, “Inception recurrent convolutional neural network for object recognition,” Machine Vision and Applications, vol. 32, no. 1, p. 28, Jan. 2021, doi: 10.1007/s00138-020-01157-3.

[36] Jahandad, S. M. Sam, K. Kamardin, N. N. Amir Sjarif, and N. Mohamed, “Offline Signature Verification using Deep Learning Convolutional Neural Network (CNN) Architectures GoogLeNet Inception-v1 and Inception-v3,” Procedia Computer Science, vol. 161, pp. 475–483, 2019, doi: 10.1016/j.procs.2019.11.147.

[37] X. Wang et al., “A Recognition Method of Ancient Architectures Based on the Improved Inception V3 Model,” Symmetry, vol. 14, no. 12, p. 2679, Dec. 2022, doi: 10.3390/sym14122679.