iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 Issue: IX Month of publication: September 2025

DOI: https://doi.org/10.22214/ijraset.2025.74113

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue IX Sep 2025- Available at www.ijraset.com

Classification of Plant Leaf Diseases Using ResNet18 Enhanced with Inception and Capsule Network

Vemuluri Divya[1], Durga Ganga Rao Kola[2]
[1]PG Student, [2]Assistant Professor, Department of Electronics and Communication Engineering, UCEK (JNTU), Kakinada, India[1,2]

Abstract: Accurate and early detection of plant leaf diseases is crucial for ensuring crop health and improving agricultural productivity. This work proposes a hybrid deep learning model that combines ResNet18, Inception blocks, and fully connected Capsule layers to classify leaf images of apple, grape, and corn plants into healthy or diseased categories. ResNet18 is used as the backbone for deep feature extraction, while Inception modules enhance the network's ability to capture multi-scale patterns. Capsule layers are employed at the final stage to retain spatial relationships and pose information, improving the model's ability to recognize complex disease features. The model is trained and evaluated using images from the PlantVillage dataset, with separate configurations for each crop. The proposed model achieved validation accuracies of 99.84% for apple, 100% for grape, and 97.27% for corn. Performance is further assessed using precision, recall, and F1-score, and compared against a baseline ResNet18 model. The results demonstrate that the proposed architecture significantly improves classification accuracy and feature understanding, making it a strong candidate for real-world agricultural disease monitoring systems.
Keywords: Plant Leaf Disease, Deep Learning, ResNet18, Inception Block, Capsule Network

I. INTRODUCTION

Agriculture plays a vital role in sustaining the global economy, especially in countries like India, where a large portion of the population depends on farming for their livelihood. However, the agricultural sector is often challenged by a wide range of plant diseases that can drastically reduce crop yield, quality, and farmer income. Plant leaf diseases, in particular, are among the most common and visually detectable symptoms of plant health issues. Timely and accurate identification of these diseases is essential to control their spread and minimize the damage. Traditional methods of disease detection rely heavily on manual inspection by trained experts or agricultural extension officers. While expert observation is often effective, it is not scalable or reliable across large farms, varied geographic regions, or real-time monitoring needs. The demand for faster, scalable, and automated plant disease detection has led researchers to explore artificial intelligence, especially deep learning, as a powerful alternative.

In recent years, deep learning models particularly convolutional neural networks (CNNs) have demonstrated exceptional performance in computer vision tasks, including image classification, object detection, and segmentation. These models learn hierarchical feature representations from data and have the capacity to capture complex spatial patterns directly from raw pixel values. CNN-based models like VGG, AlexNet, GoogLeNet, and ResNet have been widely adopted in various agricultural applications, including plant disease detection. However, despite their success, traditional CNNs have limitations. They often struggle to generalize across diseases with similar visual features, and they may fail to preserve spatial hierarchies and orientation relationships between features. Additionally, many standard architectures rely heavily on fully connected layers at the end of the network, which flatten feature maps and discard spatial structure, limiting their effectiveness in certain classification tasks.

To address these challenges, researchers have begun exploring hybrid architectures that combine the strengths of multiple network types to improve learning capability and robustness. This work proposes a hybrid deep learning model that integrates ResNet18, Inception blocks, and Capsule Network layers to classify plant leaf diseases in apple, grape, and corn plants. The proposed architecture is designed to overcome the shortcomings of traditional CNNs by improving both multi-scale feature learning and spatial relationship retention. ResNet18 is used as the core backbone of the model due to its ability to train deeper networks using identity-based residual connections, which help avoid the vanishing gradient problem. Inception blocks are integrated into the network to enhance its ability to capture features at multiple scales using parallel convolutions with different kernel sizes. Finally, fully connected Capsule layers are introduced toward the end of the model to retain the orientation of learned features, enabling more nuanced classification even when symptoms are rotated or vary in position.

The selection of ResNet18 as the base model is driven by its relatively lightweight design and proven performance in classification tasks. Unlike deeper versions of ResNet such as ResNet50 or ResNet101, ResNet18 offers a balance between computational efficiency and representational power. Its residual connections allow the network to reuse activations from earlier layers, facilitating better gradient flow and faster convergence. However, ResNet-18 alone may not be sufficient to capture the wide variation in texture, color, and pattern seen in different plant diseases. This is where Inception modules prove valuable. Originally introduced in GoogLeNet, Inception blocks apply 1×1, 3×3, and 5×5 convolutions in parallel, allowing the network to learn both local and global patterns without increasing computational cost dramatically. By incorporating Inception modules after the ResNet feature extractor, the model becomes capable of learning a richer and more diverse set of visual features.

The final component the Capsule Network is inspired by the need to retain spatial relationships between features. While standard dense layers reduce high-dimensional feature maps to a single vector, Capsule layers preserve part-to-whole relationships by representing features as vectors rather than scalars. These vectors can capture properties like orientation, position, and scale, which are especially useful when classifying leaf diseases that appear at different angles or in various shapes. Fully connected Capsule layers are used in this model instead of the original dynamic routing version proposed by Hinton, offering a more efficient implementation while still preserving spatial hierarchies.

An important advantage of the proposed model is its ability to generalize across multiple crops and disease categories using a unified architecture. By applying the same model design to apple, grape, and corn leaves, this work demonstrates that a single framework can be reused across diverse datasets without major architectural changes. This makes the model highly adaptable and scalable for practical use in agricultural diagnostics. Whether deployed in mobile applications, drone-based farm monitoring systems, or edge devices, the ability to handle various crops and disease types with consistent accuracy is critical. The model’s modular structure also allows for easy extension additional crops or new disease classes can be incorporated by retraining or fine-tuning on expanded datasets. This flexibility supports the broader goal of building intelligent agricultural systems that can function reliably across real-world farming conditions and diverse geographical regions.

Overall, this work aims to develop a generalizable and efficient system for plant leaf disease detection that can be adapted across multiple crop types. By combining multiple deep learning components into a single architecture, the proposed model leverages the strengths of each to achieve higher accuracy and better interpretability. The hybrid approach demonstrates the potential of deep learning not only as a classification tool but also as a step toward building intelligent decision support systems in agriculture. With increasing availability of labeled data and computing resources, such models can eventually be deployed in real-world settings, including mobile applications, drone-based monitoring, and automated farm systems.

II. LITERATURE REVIEW

Hosny et al. [1] introduced a lightweight deep CNN model for multi-class classification of plant leaf diseases by combining deep learned features with handcrafted LBP features. The model was trained and tested on three public datasets namely Apple, Tomato, and Grape leaf images while data augmentation techniques like rotation, flipping, and scaling were used to balance the datasets. The architecture consisted of three convolution layers followed by max-pooling layers, four fully connected layers, and ReLU and dropout regularization to avoid overfitting. Using a CNN, high-level features were extracted whereas LBP captured local texture information, and these were concatenated at the flatten layer, from where softmax function was used for classification. Accuracy, precision, recall, F1-score, confusion matrix, and AUC-ROC curve were used as performance metrics. The proposed method validated with the Apple dataset at an accuracy of 99%, the Tomato dataset at an accuracy of 96.6%, and the Grape dataset at an accuracy of 98.5%, whereas the test accuracies were 98.8%, 96.5%, and 98.3%, respectively.

Madhurya and Jubilson [2] developed a highly optimized deep learning system called YR2S for faster detection and classification of plant leaf diseases. Contrast enhancement was accomplished via the CLAHE algorithm; feature extraction was operated under Pyramid Channel-based Feature Attention Network (PCFAN), while detection of disease-affected regions was handled by the YOLOv7 object detection model. These classified regions were then forwarded toward ShuffleNetV2 for the categorization of diseases, and hyperparameter tuning of the ShuffleNetV2 model was conducted using the Enhanced Rat Swarm Optimization (ERSO) algorithm to increase accuracy and cut down the computational time in the process. The segmentation of diseased spots is performed by employing the Red Fox Optimization (RFO) Algorithm. The model was validated using a set of images consisting of both healthy and diseased plant leaves. Experimental results shows that this method attained 99.69% accuracy in classification and outperformed well-known approaches such as OMNCNN, DenseNet121, AlexNet, VGG16+ResNet50, and SVM in terms of classification accuracy and computational speed.

Moupojou et al. [3] described FieldPlant, a truly versatile database set up for deep learning plant disease detection and classification in real agricultural conditions. The dataset construction involved the collection of 5,170 field images from plantations in Cameroon, showing diseased cassava, corn, and tomato crops. The manual annotation of 8,629 leaves representing 27 disease classes was supervised by expert plant pathologists to maximize reliability in labeling. In order to test FieldPlant's usefulness, several deep learning models-MobileNet, VGG16, InceptionV3, and InceptionResNetV2-were benchmarked in classification using both raw and cropped leaf images, and results were compared with PlantVillage and PlantDoc, two of the more used datasets in the field. Classification accuracy dropped drastically when models trained on PlantVillage were evaluated on FieldPlant because of background complexity and image structure. However, in both classification and detection tasks, FieldPlant proved better than PlantDoc, demonstrating its application to real-world plant disease detection. The results indicated MobileNet achieved relatively better accuracy among the models tested, highlighting the dataset's potential for advancing deep learning-based plant disease detection under practical field conditions.

Balafas et al. [4] conducted a review of 79 articles associated with machine learning and deep learning methods for detecting and classifying plant diseases in precision agriculture. The articles were separated into classification methods and detection methods, and a compiled framework for evaluation and comparison was proposed. The PlantDoc dataset was used for in-depth analysis, in which 18 classification and 5 detection models were reviewed. The best performing model for classification was ResNet50, with a best accuracy of 97.1%, and MobileNetv2 with a best accuracy of 95.2%. The best performing detection model was YOLOv5, achieving a mean average precision (mAP) of 79.3%. The authors recognized the need to use datasets obtained from actual field cases, and to develop and adopt efficient models that can be tested in an agricultural environment.

Benfenati et al. [5] worked on unsupervised deep learning methods for detecting powdery mildew disease in cucumber leaves without using any manual image labeling. They implemented two models: one clustering convolutional autoencoder (Clu-AE) and one anomaly detection (Ano-AE) model based on residual autoencoders. Multispectral images were analyzed with RGB and NIR bands. The Clu-AE gave unsatisfactory results, as clustering was more influenced by leaf shape rather than by disease severity. The Ano-AE, however, offered much better results with classification accuracy up to 90.35% and an AUC of 0.91, using feature reconstruction error as the anomaly score. This study showed that the anomaly detection approach was more useful for unsupervised identification of plant diseases and provided an attractive alternative to reduce dependency on manually labeled datasets.

Hama et al. [6] proposed a houseplant leaf classification system based on deep learning with an improved ResNet-50 model. In total, a new dataset of 2,500 images across ten houseplant species was constructed, rather a new dataset was constructed and expanded with five data augmentation methods to facilitate effective training of the model and minimize overfitting. The improved ResNet-50 architecture included hyperparameter tuning of the model, as well as selective layer freezing to reduce complexity and improve performance. Furthermore, comparative experiments were conducted with both the original ResNet-50 and MobileNetv2. The improved model achieved the highest classification accuracy of 99.00%, precision 99.03%, recall 99.00%, and an F1 score of 99.01% on the augmented dataset across all plants. On the non-augmented dataset, it attained an accuracy of 98.60%, establishing its robustness and efficiency for houseplant species recognition tasks.

Babu et al. [7] for the identification of rice plant disease in which CNN structures ResNet-101, InceptionV4, VGGNet-16, and DenseNet-121 were employed. The aim of the study was to detect four common rice plant diseases that included leaf smut, bacterial leaf blight, sheath blight, and brown spot. The data for the study were captured from the PlantVillage dataset for model development and testing. All the models were specifically tuned and analysed based on training as well as validation accuracy and loss. DenseNet-121 performed the best, attaining a final validation accuracy of 99.17%. ResNet-101 achieved 99.84% accuracy, Inception V4 - 99.62% and VGGNet-16 attained 90.81%. DenseNet performed above rest of the models in accuracy and efficiency as well as was very effective for the accurate and efficient detection of rice plant diseases utilizing deep learning techniques.

Maurya et al. [8] designed a lightweight deep learning architecture in RAI-Net, combining a fine-tuned ResNet18 model with a channel attention mechanism and an Inception module for tomato plant diseases classification. The specific aims of the model were to improve multiscale feature extraction and attentiveness to disease-specific portions of the image. The input images were preprocessed, resized, and augmented for better model generalization. RAI-Net was tested by using a dataset consisting of 22,930 images from 10 class categories, two of which were the healthy and diseased leaf. RAI-Net achieved a classification accuracy of 97.88 percent. In addition to RAI-Net, three other baseline models were referenced: ResNet18, SE- ResNet, and InceptionV3. RAI-Net outperformed the baseline models and the ablation studies demonstrated that channel attention entries and Inception modules contributed positively to the overall performance. The grad-CAM visualizations concerning the tomato leaf images showed the RAI-Net became aware of the image portions which were critical to tomato disease, and demonstrated this attentiveness to disease portions.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue IX Sep 2025- Available at www.ijraset.com

Noon et al. [9] developed a modified YOLOX-based deep learning model to identify multiple co-occurring cotton plant diseases and their different severity stages on a single leaf. The model added an altered Spatial Pyramid Pooling (SPP) block and used α-IoU loss to improve the bounding box regression. Data augmentation was increased through multiple mosaic and mixup methods to increase dataset diversity. The model was evaluated on a custom dataset of 1,112 annotated images of cotton leaves that shared symptoms of overlapping cotton leaf curl and sooty mold. The improved YOLOX model maintain a 73.13% mean average precision (mAP) score on training data and 72.31% on test data performance, beating baseline YOLOX and other models such as YOLOv4, YOLOv5, and EfficientDet in overall accuracy when identifying a more complicated detection task.

Theerthagiri et al. [10] built a deep learning model based on SqueezeNet architecture to classify maize leaf diseases, such as blight, common rust, grey leaf spot, and healthy leaves. Using SMOTE for class balancing, the model used extensive data augmentation to improve generalization. To measure performance, accuracy, precision, recall, and F1-score were calculated and benchmarked against VGG16, ResNet34, and ResNet50. In terms of accuracy, the proposed SqueezeNet model provided the highest classification accuracy (97.00%) with precision (98.00%), recall (95.00%), and F1-score (96.00%). The proposed SqueezeNet model decreased the mean square error by 4–11% compared to the other models and also maintained fewer parameters, so it is also lightweight and fast (real-time diagnosis) and provides success for using DNN to detect diseases in an agricultural field plant.

III. PROPOSED MODEL

The proposed model aims to improve the accuracy and robustness of plant leaf disease classification by combining three key deep learning components: ResNet18, Inception blocks, and a fully connected Capsule Network, followed by specialized classification layers. Each component plays a distinct role: ResNet18 serves as a strong backbone for hierarchical feature extraction, Inception blocks enhance the model's ability to learn from multi-scale visual patterns, and Capsule layers help retain spatial and structural relationships that traditional CNNs often fails. Together, this combination creates a unified framework capable of handling complex and subtle disease variations in leaf images across multiple crop types. The Proposed model is shown in Figure 1.

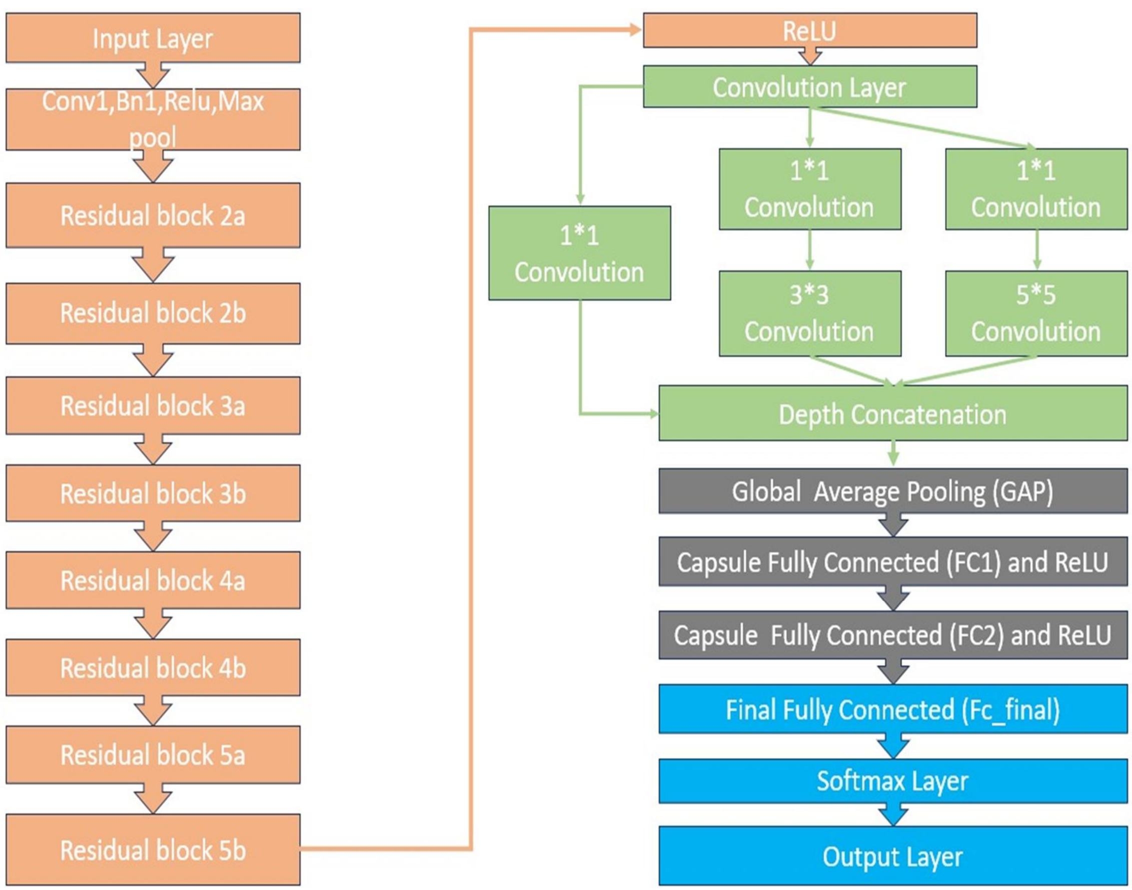


Figure 1. Proposed Model

A. ResNet18 Backbone

At the core of the proposed model lies ResNet18, a deep convolutional neural network known for its residual learning framework. Unlike traditional CNNs, which struggle to maintain performance as depth increases, ResNet solves the vanishing gradient problem through identity-based skip connections. These shortcuts allow gradients to flow directly through earlier layers, ensuring effective training of deeper networks.

ResNet18 consists of 18 layers with learnable weights, including a series of convolutional layers, batch normalization, ReLU activations, and residual blocks.

Each residual block includes two convolutional layers and a shortcut path that bypasses them. The output of each block is the element-wise sum of the convolutional path and the identity path, allowing the network to learn both transformed and untransformed feature representations.

In the proposed model, ResNet18 is used up to the final residual block (layer 5b). This allows the network to extract deep hierarchical features from the input leaf images, capturing edges, textures, color variations, and early shape patterns. These features form the foundation upon which the rest of the model builds, enabling more advanced modules like Inception and Capsule to focus on higher-order relationships.

B. Inception Block Integration

While ResNet is excellent at learning deep features, it tends to use fixed kernel sizes within each layer, which can limit its sensitivity to visual details occurring at different spatial scales. To address this limitation, the proposed model integrates an Inception block after the final ResNet18 layer.

An Inception block is a modular design introduced in GoogLeNet that applies multiple convolutional filters in parallel to the same input. These typically include 1×1, 3×3, and 5×5 convolutions. The 1×1 filters act as dimensionality reducers, helping control the number of parameters and computational cost. The 3×3 and 5×5 filters allow the network to capture both medium and large patterns in the image.

This parallel structure enables the Inception block to learn multi-scale features, which are particularly important for plant disease detection, where lesions, discoloration, or fungus may appear in very different shapes and sizes.

In the proposed model, the output from the last ResNet block is fed into an Inception module consisting of three parallel branches:

- One with a 1×1 convolution
- One with a 1×1 followed by a 3×3 convolution
- One with a 1×1 followed by a 5×5 convolution

All branches are followed by batch normalization and ReLU activation. The outputs are concatenated along the depth dimension using depth concatenation, forming a rich, multi-scale representation of the input feature map. This output serves as the input to the Capsule Network module.

C. Capsule Network Layer

Conventional CNNs typically flatten the learned feature maps before sending them to fully connected layers. This process destroys spatial relationships between features an important loss, especially when classifying images with complex arrangements, such as leaves with diseases that vary in orientation, shape, or distribution.

To overcome this, the proposed model replaces traditional dense layers with a fully connected Capsule layer. Capsule Networks, introduced by Hinton et al., represent features as vectors rather than scalars. Each vector encodes the presence of a feature along with its spatial properties such as pose, orientation, and scale. This allows the model to understand not just what features exist but how they are positioned and related to one another.

In this model, the Capsule layer is implemented in two stages:

The first stage (FC1) is a fully connected layer with ReLU activation, transforming the multi-scale output from the Inception block into vector capsules.

The second stage (FC2) further refines these capsules and prepares them for final classification.

Although dynamic routing between capsules (as used in original CapsuleNet) can be computationally heavy, here a fully connected capsule approach is used to maintain a balance between performance and efficiency. This design retains important spatial hierarchies while keeping the model trainable on standard hardware.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue IX Sep 2025- Available at www.ijraset.com

D. Classification Layers

After the Capsule layers, the refined vector representation is fed into the final classification module. This includes three main layers:

1) Final Fully Connected Layer (Fc_final): This layer transforms the high-dimensional capsule output into a vector whose size equals the number of disease classes in the dataset (including a class for healthy leaves).
2) Softmax Layer: It converts the raw class scores from Fc_final into probabilities, ensuring that all outputs sum to one.
3) Output Layer: This layer selects the class with the highest probability as the final prediction.

This classification stack allows the model to provide clear and confident predictions, backed by a robust representation of both low-level visual cues and high-level spatial relationships.

E. Summary of the Architecture Flow

The input image is resized and passed through the ResNet18 backbone for deep feature extraction. Output from ResNet is passed to an Inception block to extract multi-scale features. These features are sent to a Capsule layer that preserves spatial structure and relationships. The capsule output is passed through classification layers, ending in a softmax output. The model predicts the disease class based on the highest probability.

This architecture is designed to perform well across different crops and disease patterns by combining the depth of ResNet, the multi-scale awareness of Inception, and the spatial encoding of Capsule layers

IV. EXPERMENTATION

A. Datasets

To train and evaluate the proposed deep learning model for plant leaf disease classification, this work uses image datasets of three widely cultivated crops: apple, grape, and corn. These crops are not only economically important but also vulnerable to a range of foliar diseases that can significantly affect yield and quality. The datasets used contain labeled images of both healthy and diseased leaves, covering several common diseases for each crop. All images are sourced from the PlantVillage dataset, which provides high-quality, standardized leaf images captured under controlled conditions. This ensures consistency across samples and helps the model focus on identifying disease-related patterns such as color changes, spots, blight, and lesions

1) Apple Dataset

The apple leaf dataset contains a total of 3,171 images and these are divided into four categories: Apple Scab, Black Rot, Cedar Apple Rust, and Healthy. Each disease class exhibits distinct visual symptoms, such as dark, velvety patches for Apple Scab, circular lesions for Black Rot, and bright orange or rust-colored spots for Cedar Apple Rust. The images were captured under uniform lighting and plain backgrounds, typically at a resolution of 256×256 pixels, to reduce background noise and highlight leaf features. This dataset provides sufficient variety in texture, color, and disease severity, making it suitable for training the proposed model to learn fine-grained differences. Sample images from the apple dataset are shown in Figure 2.

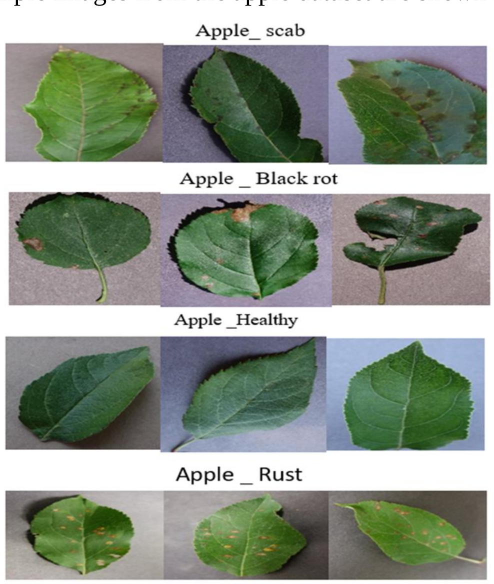


Figure 2. Apple Dataset Images

2) Grape Dataset

The grape leaf dataset used in this work is also sourced from the PlantVillage collection and contains a total of 3,503 images. It includes four classes: Black Rot, Esca (Black Measles), Leaf Blight (Isariopsis Leaf Spot), and Healthy. The disease symptoms vary across classes, with Black Rot showing dark, circular lesions, Esca presenting as interveinal discoloration and necrotic patches, and Leaf Blight appearing as elongated brown spots along the leaf veins. Images were captured under controlled conditions with plain backgrounds and consistent lighting, typically at 256×256 resolution, allowing the model to focus on disease-specific patterns without background interference. The diversity in symptoms and textures across classes helps in training a robust classifier. Sample images from the grape dataset are shown in Figure 3.

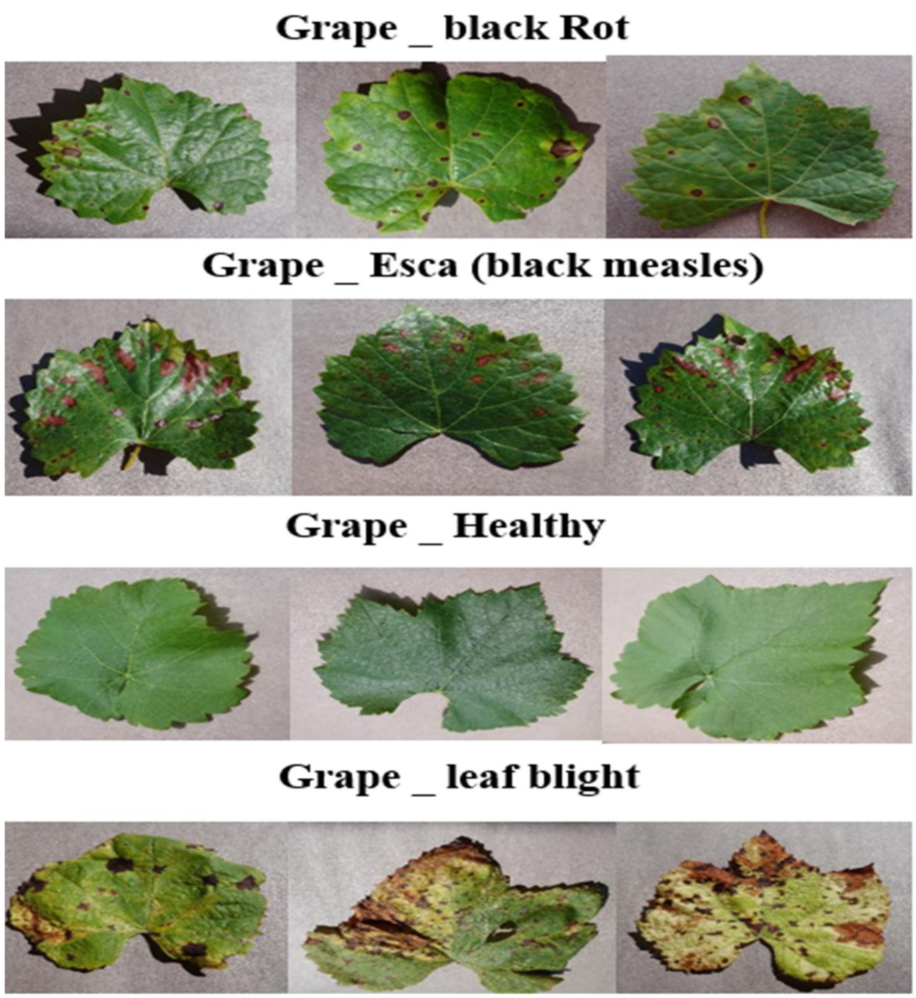


Figure 3. Grape Dataset Images

3) Corn Dataset

The corn leaf dataset used in this work consists of 3,852 images taken from the PlantVillage collection. It is organized into four classes: Gray Leaf Spot (Cercospora Leaf Spot), Common Rust, Northern Leaf Blight, and Healthy. Each disease presents distinct visual features Gray Leaf Spot appears as narrow, rectangular lesions between veins, Common Rust shows reddish brown pustules, and Northern Leaf Blight is characterized by long, grayish lesions along the leaf surface. The images were captured in controlled environments with consistent lighting and plain backgrounds, typically resized to 256×256 pixels to standardize input for training. This dataset provides clear distinctions between healthy and diseased leaves, enabling the model to learn subtle visual cues effectively. Sample images from the corn dataset are shown in Figure 4.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue IX Sep 2025- Available at www.ijraset.com

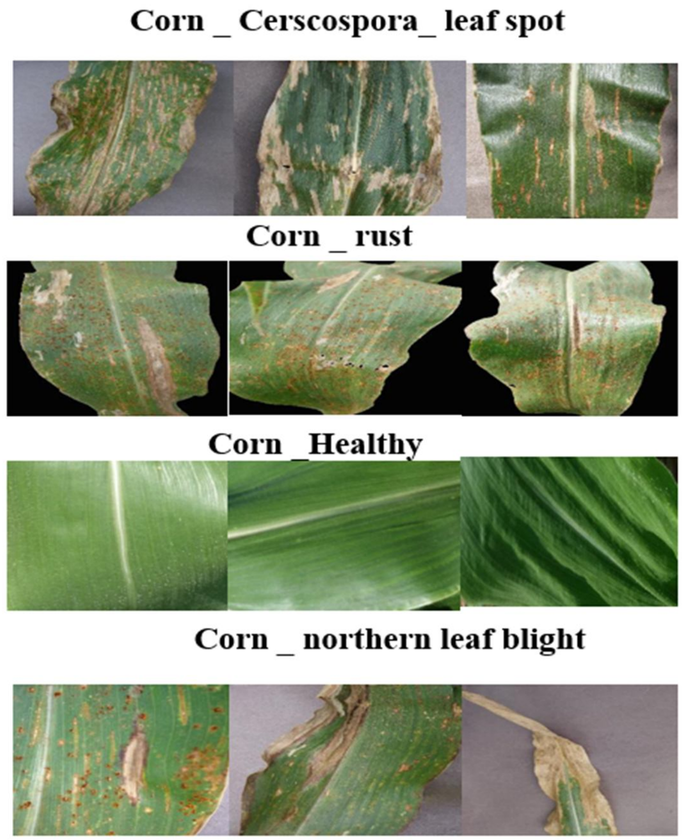


Figure 4. Corn Dataset Images

B. Performance Evaluation

To assess the effectiveness of the proposed leaf disease classification model, several evaluation metrics are used. These include accuracy, precision, recall, and F1-score, which provide a detailed understanding of the model's ability to correctly identify healthy and diseased leaves. These metrics are calculated using the values obtained from the confusion matrix—specifically, true positives, false positives, true negatives, and false negatives. Evaluating the model using multiple metrics ensures a balanced view of its performance, especially in cases where class distribution is not perfectly balanced.

1) *Accuracy* measures the overall correctness of the model by calculating the proportion of correctly classified samples out of the total number of samples. It is defined as:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

where *TP* is true positives, *TN* is true negatives, *FP* is false positives, and *FN* is false negatives.

2) *Precision* evaluates how many of the predicted positive cases are correct. It is especially useful in applications where false positives are costly. The formula for precision is:

$$\text{Precision} = \frac{TP}{TP+FP}$$

3) *Recall* also known as sensitivity or true Precision positive rate, measures the model's ability to correctly identify actual positive cases. It is defined as:

$$\text{Recall} = \frac{TP}{TP + FN}$$

4) *F1-score* provides a harmonic mean of precision and recall, offering a single metric that balances both false positives and false negatives. It is particularly useful when dealing with imbalanced datasets. It is defined as:

$$\text{F1-score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

C. Software Tool

The proposed model is implemented and evaluated using MATLAB 2025a, with support from the Deep Learning Toolbox. This toolbox enabled the design and training of the custom architecture, including integration of specialized components such as activation functions and attention mechanisms. The training process, performance evaluation, and result visualization are carried out within the MATLAB environment, providing a streamlined and consistent workflow for experimentation with the proposed approach.

D. Training and Validation Data

In this work, the dataset is divided into training and validation sets using an 80:20 split, applied separately to each class to maintain balanced representation. The model was trained for 6 epochs on the apple and grape datasets, and for 10 epochs on the corn dataset. A batch size of 32 was used consistently across all experiments. The learning rate was set to 0.001, allowing stable and gradual updates to the model weights during training. All input images were resized to match the input dimensions required by the proposed model architecture. No preprocessing or augmentation techniques were applied.

V. RESULTS

A. Performance on Apple Dataset

For the apple dataset, the training and validation accuracy and loss curves are shown in Figure 5 indicate smooth convergence and no signs of overfitting. It is observed that, the proposed model achieved a validation accuracy of 99.84%, demonstrating its ability to reliably differentiate between healthy leaves and diseases such as Apple Scab, Black Rot, and Cedar Apple Rust. The confusion matrix is shown in Figure 6 which indicates that most predictions are correct, with very few misclassifications between similar disease classes. The model was particularly strong in detecting Cedar Apple Rust, likely due to its visually distinct appearance.

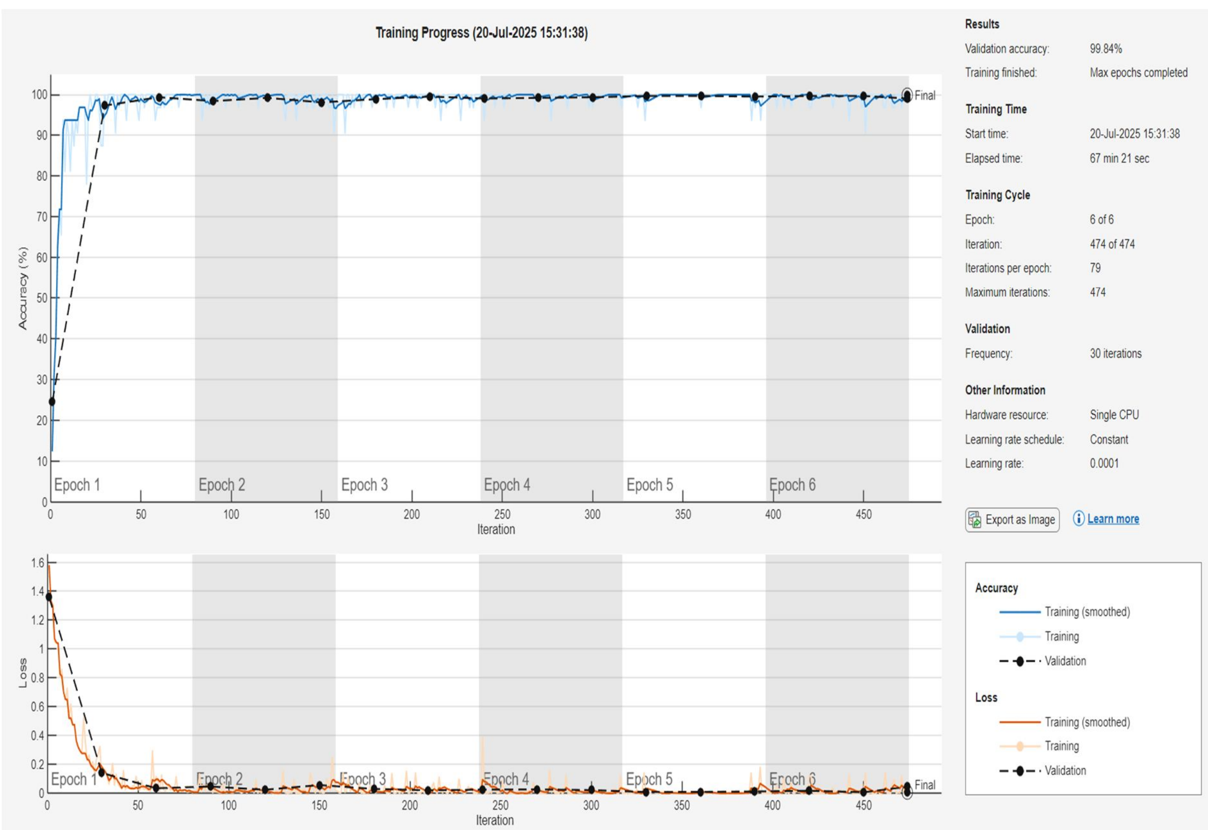


Figure 5. Training and Validation Accuracy and Loss Curves for the Apple Dataset using Proposed Model

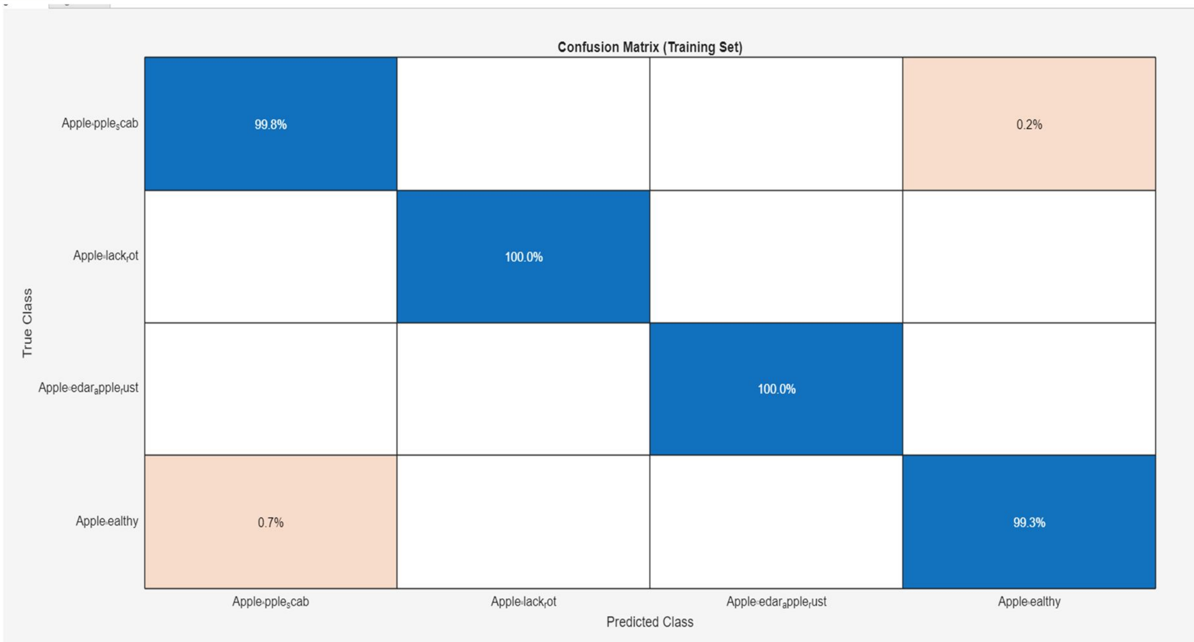


Figure 6. Confusion Matrix of the Proposed Model on the Apple Dataset

B. Performance on Grape Dataset

For the grape leaf dataset, the proposed model achieved a validation accuracy of 100%, confirming its effectiveness in identifying various conditions such as Black Rot, Esca (Black Measles), Leaf Blight, and healthy grape leaves. The accuracy and loss graphs, shown in Figure 7, demonstrate stable and consistent learning throughout training. The accuracy curve rises smoothly with epochs, and the validation curve closely follows the training curve, indicating that the model maintains performance without overfitting. The loss curves show a clear downward trend, further validating the model's effective convergence. The confusion matrix in Figure 8 shows strong diagonal dominance, reflecting high classification accuracy across all classes and confirming that the model has successfully captured the relevant visual patterns for grape leaf disease classification.

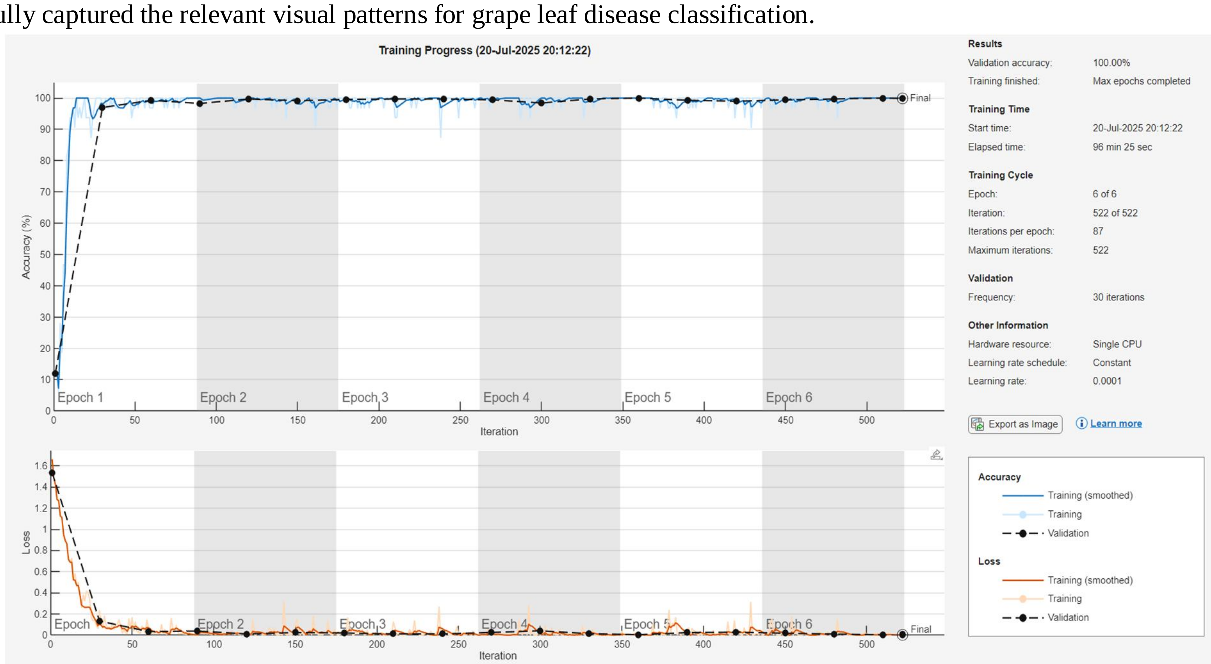


Figure 7. Training and Validation Accuracy and Loss Curves for the grape Dataset using Proposed Model

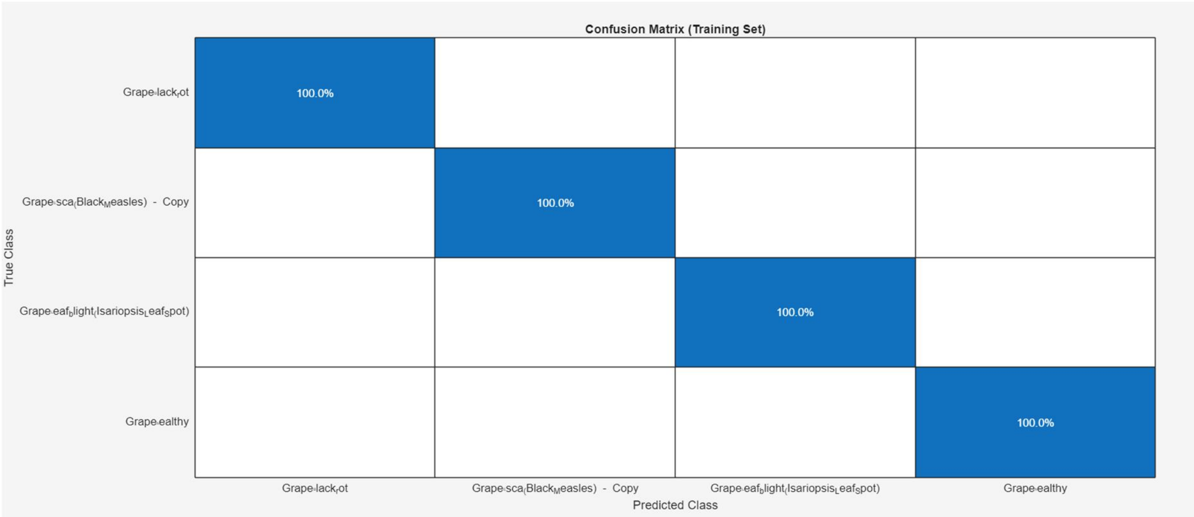


Figure 8. Confusion Matrix of the Proposed Model on the Grape Dataset

C. Performance on Corn Dataset

For the Corn leaf dataset, the model attained a validation accuracy of 97.27%, showing high performance in recognizing Gray Leaf Spot, Common Rust, Northern Leaf Blight, and healthy leaves. The training and validation accuracy curves, displayed in Figure 9, show consistent improvement during the training phase, with validation accuracy closely tracking the training accuracy. The loss curves show a steady decline, with minimal fluctuations, indicating that the model is learning effectively across epochs. As illustrated in the confusion matrix in Figure 10, most of the predictions fall along the diagonal, which confirms that the model accurately classifies corn leaf images across all categories. This result highlights the model's ability to extract and learn fine-grained disease features present in corn leaves.

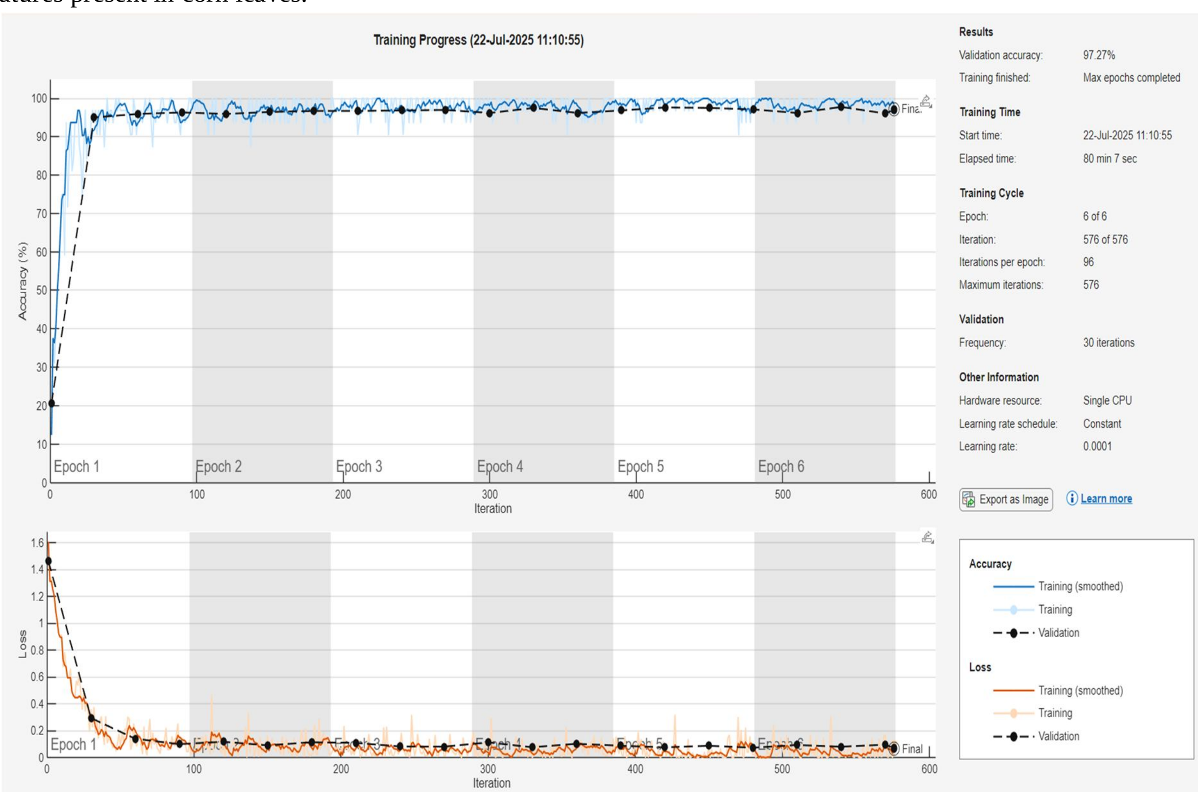


Figure 9. Training and Validation Accuracy and Loss Curves for the Corn Dataset using Proposed Model

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue IX Sep 2025- Available at www.ijraset.com

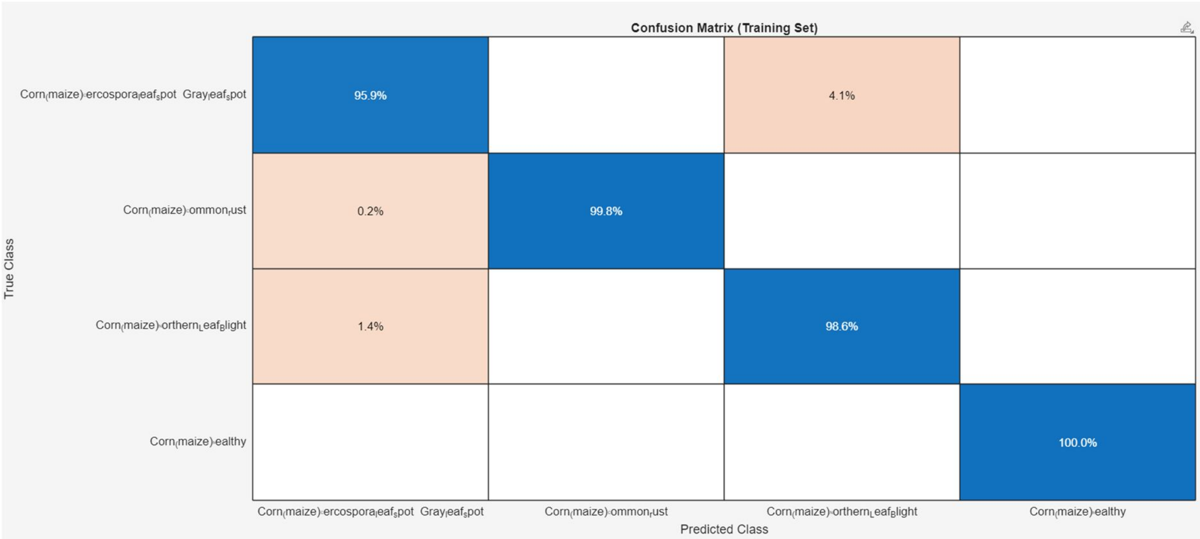


Figure 10. Confusion Matrix of the Proposed Model on the Corn Dataset

D. Comparative Analysis

To assess the effectiveness of the proposed architecture, a comparative study is conducted using a standard ResNet-18 model as the baseline. The proposed model enhances ResNet-18 by integrating Inception blocks to improve multi-scale feature extraction and a fully connected Capsule layer to better capture spatial relationships and pose information within feature maps. Both models are trained under identical conditions, with the same datasets, batch size, learning rate, and training-validation split, to ensure a fair comparison.

The first comparison focuses on classification accuracy across all three datasets apple, grape, and corn. As shown in Table 1, the proposed model consistently outperforms the baseline ResNet18, demonstrating improved feature learning and stronger generalization, particularly on datasets with more complex disease patterns. The second table presents a detailed breakdown of the proposed model's performance using precision, recall, and F1-score. These metrics provide a more complete view of how well the model handles class-specific predictions, especially when class distributions are not perfectly balanced.

Table 1: Accuracy Comparison: ResNet-18 vs Proposed Model (ResNet-18 + Inception + Capsule)

Datasets	ResNet18 Accuracy (%)	Proposed Model Accuracy (%)
Apple	96.00%	99.84%
Grape	97.10%	100.00%
Corn	95.00%	97.27%

Table 2: Performance Metrics of the Proposed Model

Datasets	Precision (%)	Recall (%)	F1-Score (%)
Apple	100.00%	100.00%	100.00%
Grape	100.00%	100.00%	100.00%
Corn	96.00%	97.00%	96.00%

V. CONCLUSION

The proposed ResNet18 model enhanced with Inception blocks and a fully connected Capsule Network achieved good results for plant leaf disease classification in apple, grape, and corn datasets. By integrating residual learning, multi-scale feature extraction, and spatially aware capsule encoding, the model effectively captured disease-specific patterns and obtained higher accuracy than the baseline ResNet18. Experimental evaluation showed that the proposed model achieved 99.84% accuracy for apple, 100% for grape, and 97.27% for corn, outperforming the standard ResNet18 across all metrics. In addition, the precision, recall, and F1-scores are consistently above 93%, confirming the robustness of the approach.

These outcomes highlight the effectiveness of combining deep learning architectures for reliable agricultural disease detection, supporting timely intervention to safeguard crop yield and minimize economic losses. For future work, the model can be extended to larger and more diverse datasets collected under real field conditions, with further improvements such as attention-based feature refinement, optimized capsule routing, and lightweight deployment on mobile or edge devices to enable real-time disease detection for precision agriculture.

REFERENCES

[1] K. M. Hosny, W. M. El-Hady, F. M. Samy, E. Vrochidou, and G. A. Papakostas, "Multi-Class Classification of Plant Leaf Diseases Using Feature Fusion of Deep Convolutional Neural Network and Local Binary Pattern," IEEE Access, vol. 11, pp. 62307–62317, 2023, doi: 10.1109/ACCESS.2023.3286730. [Online]. Available: https://doi.org/10.1109/ACCESS.2023.3286730.

[2] C. Madhurya and E. A. Jubilson, "YR2S: Efficient Deep Learning Technique for Detecting and Classifying Plant Leaf Diseases," IEEE Access, vol. 12, pp. 3789–3804, 2024, doi: 10.1109/ACCESS.2023.3343450.

[3] E. Moupojou, A. Tagne, F. Retraint, A. Tadonkemwa, D. Wilfried, H. Talamo, and M. Nkenlifack, "FieldPlant: A Dataset of Field Plant Images for Plant Disease Detection and Classification With Deep Learning," IEEE Access, vol. 11, pp. 35398–35410, 2023, doi: 10.1109/ACCESS.2023.3263042 Paper Link: https://doi.org/10.1109/ACCESS.2023.3263042.

[4] V. Balafas, E. Karantoumanis, M. Louta, and N. Ploskas, "Machine Learning and Deep Learning for Plant Disease Classification and Detection," IEEE Access, vol. 11, pp. 114352–114376, Oct. 2023, doi: 10.1109/ACCESS.2023.3324722. [Online]. Available: https://doi.org/10.1109/ACCESS.2023.3324722.

[5] A. Benfenati, P. Causin, R. Oberti, and G. Stefanello, "Unsupervised deep learning techniques for automatic detection of plant diseases: reducing the need of manual labelling of plant images," Journal of Mathematics in Industry, vol. 13, no. 5, 2023. doi: 10.1186/s13362-023-00133-6. [Online]. Available: https://doi.org/10.1186/s13362-023-00133-6.

[6] H. M. Hama, T. Sh. Abdulsamad, and S. M. Omer, "Houseplant leaf classification system based on deep learning algorithms," Journal of Electrical Systems and Information Technology, vol. 11, no. 18, pp. 1–15, Feb. 2024, doi: 10.1186/s43067-024-00141-5. [Online]. Available: https://doi.org/10.1186/s43067-024-00141-5.

[7] S. Babu, M. Maravarman, and R. Pitchai, "Detection of Rice Plant Disease Using Deep Learning Techniques," Journal of Mobile Multimedia, vol. 18, no. 3, pp. 757–770, Jan. 2022, doi: 10.13052/jmm1550-4646.18314. [Online]. Available: https://doi.org/10.13052/jmm1550-4646.18314.

[8] R. Maurya, L. Rajput, and S. Mahapatra, "RAI-Net: Tomato Plant Disease Classification Using Residual-Attention-Inception Network," IEEE Access, vol. 13, pp. 64832–64840, Apr. 2025, doi: 10.1109/ACCESS.2025.3559804. [Online]. Available: https://doi.org/10.1109/ACCESS.2025.3559804.

[9] S. K. Noon, M. Amjad, M. A. Qureshi, and A. Mannan, "Handling Severity Levels of Multiple Co-Occurring Cotton Plant Diseases Using Improved YOLOX Model," IEEE Access, vol. 10, pp. 134811–134825, Dec. 2022, doi: 10.1109/ACCESS.2022.3232751.

[10] P. Theerthagiri, A. U. Ruby, J. G. C. Chandran, T. H. Sardar, and A. B. M. Shafeeq, "Deep SqueezeNet learning model for diagnosis and prediction of maize leaf diseases," *Journal of Big Data*, vol. 11, no. 112, 2024. doi: 10.1186/s40537-024-00972-z.

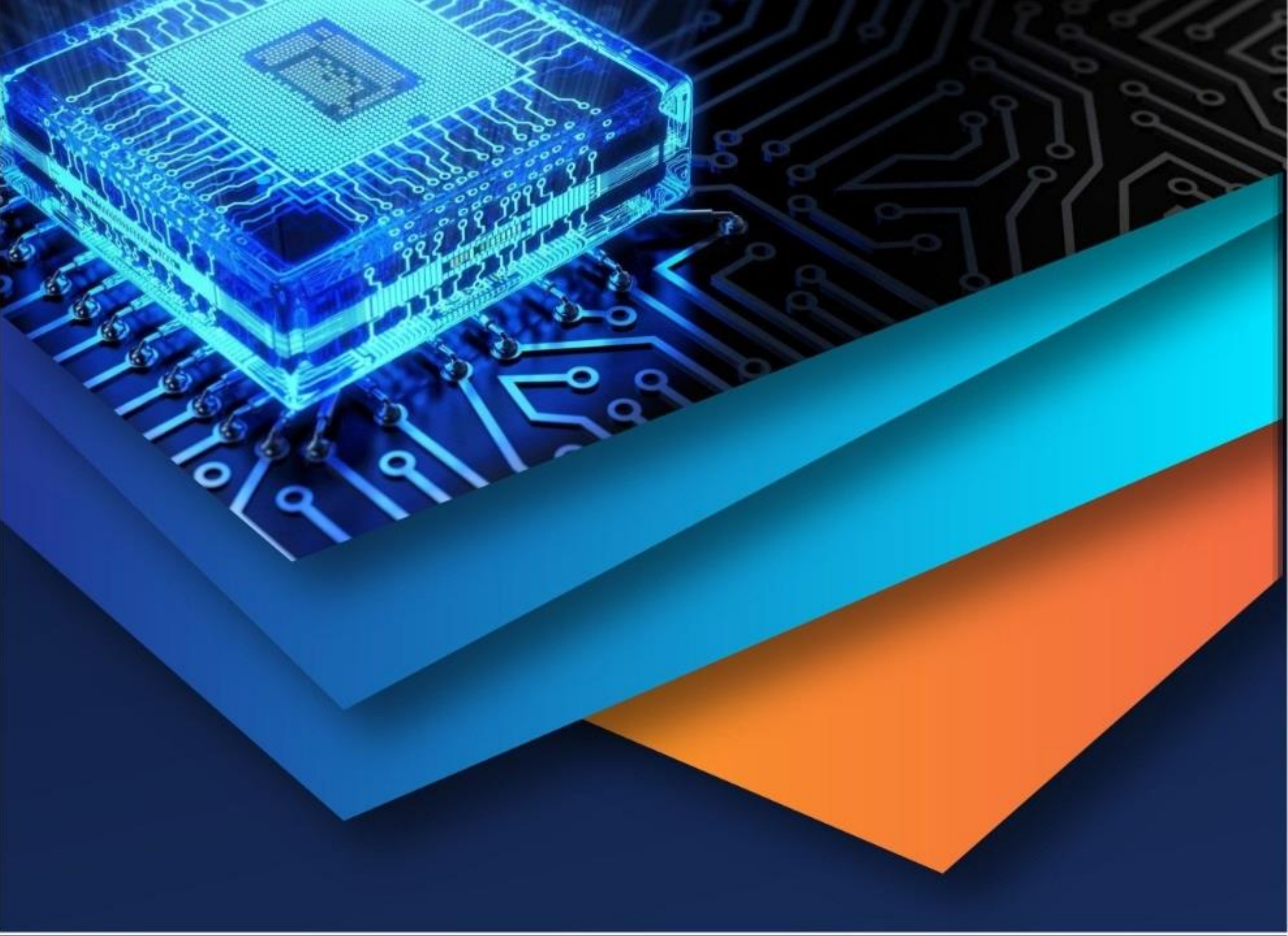
[11] S. Sabour, N. Frosst, and G. E. Hinton, "Dynamic routing between capsules," in Advances in Neural Information Processing Systems (NeurIPS), pp. 3856–3866, 2017. [Online]. Available: https://arxiv.org/abs/1710.09829.

[12] T. V. Le, L. T. Ngo, and T. D. Pham, "Detection of Co-Occurring Tomato Leaf Diseases Based on Improved YOLOX," *Sensors*, vol. 23, no. 15, p. 6841, Aug. 2023, doi: 10.3390/s23156841. [Online]. Available: https://doi.org/10.3390/s23156841.

[13] Y. Liu, Q. Wang, Y. Liu, and M. Ma, "A plant leaf disease image classification method integrating capsule network and residual network," Computational Intelligence and Neuroscience, vol. 2022, Art. no. 1759507, pp. 1–12, 2022, doi: 10.1155/2022/1759507.

[14] K. P. A. Rani and S. Gowrishankar, "Pathogen-based classification of plant diseases: A deep transfer learning approach for intelligent support systems," IEEE Access, vol. 11, pp. 64476–64493, Jun. 2023, doi: 10.1109/ACCESS.2023.3284680.

[15] J. Chaki and D. Ghosh, "Deep learning in leaf disease detection 2014–2024: A visualization-based bibliometric analysis," *IEEE Access*, vol. 12, pp. 95291–95309, Jul. 2024, doi: 10.1109/ACCESS.2024.3425897.



10.22214/IJRASET

INDEX COPERNICUS

45.98

IMPACT FACTOR:
7.129

IMPACT FACTOR:
7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)