International Information and Engineering Technology Association
Advancing the World of Information and Engineering

Ingénierie des Systèmes d'Information
Vol. 31, No. 2, February, 2026, pp. 357-369
Journal homepage: http://iieta.org/journals/isi

A Transfer Learning-Based Hybrid Deep Learning Framework for Multi-Crop Plant Leaf Disease Classification

Anton[1,2*], Supriadi Rustad[1], Guruh Fajar Shidik[1], Abdul Syukur[1]

[1] Faculty of Computer Science, Universitas Dian Nuswantoro, Semarang 50131, Indonesia
[2] Faculty of Information Technology, Universitas Nusa Mandiri, Jakarta 13620, Indonesia

Corresponding Author Email: anton@nusamandiri.ac.id

https://doi.org/10.18280/isi.310204

Received: 27 August 2025
Revised: 15 November 2025
Accepted: 18 February 2026
Available online: 28 February 2026

Keywords:
plant leaf disease detection, transfer learning, hybrid deep learning, DenseNet121, MobileNetV2, Grad-CAM++, edge computing, multi-crop classification

ABSTRACT

Early and accurate detection of plant leaf diseases is crucial for sustainable crop management, as such diseases threaten global agricultural productivity and food security. This study introduces a hybrid deep learning framework based on transfer learning for the multi-class classification of plant leaf diseases in cassava, wheat, and tomato. The proposed architecture integrates DenseNet121 and a scaled MobileNetV2 ($\alpha = 0.35$) to achieve a balance between representation power and computational efficiency. The framework incorporates attention mechanisms, compresses feature maps using 1 × 1 convolutions, and fuses them in a compact classification head using Swish activation and batch normalization. Gradient-weighted Class Activation Mapping++ (Grad-CAM++) is used for model interpretability, highlighting disease-relevant regions. Evaluated on three public datasets comprising 10,635 images across 13 disease classes and one healthy class, the model achieves perfect accuracy on cassava, approximately 95% accuracy on wheat, and up to 99.6% on tomato. Despite a compact design with ~10 million parameters, the model performs competitively and is suitable for deployment on edge devices, with inference latency under 60 ms and throughput above 20 Frame per Second (FPS).

1. INTRODUCTION

The world's population is expected to reach close to 10 billion by 2050, necessitating a 70% increase in food production over current levels, posing significant challenges for global food supply [1]. Cassava (Manihot esculenta) [2], wheat (Triticum spp.) [3], and tomato (Solanum lycopersicum) [4] are globally important food commodities, significantly impacting food security and agriculture. Cassava, a staple food for many people in sub-Saharan Africa, supports approximately 700 million people and provides essential starch for various industries, including food processing and biofuels [5]. Wheat, a major cereal crop, is crucial to diets worldwide, especially in temperate regions. Its versatility in culinary applications has encouraged widespread cultivation, making it an important contributor to the global food production system [6]. Meanwhile, tomato, a major vegetable crop, valued not only for its nutritional content but also for its economic importance in the global market, faces threats from pests and diseases that can affect yields and economic returns [3]. Therefore, continued investment in agricultural research and management strategies is crucial for the resilience of this staple food crop to biotic and abiotic stressors [7, 8].

Cassava, wheat, and tomatoes are susceptible to various leaf diseases that significantly reduce productivity. Examples include Cassava Mosaic Disease and Cassava Brown Streak Disease in cassava [9, 10], rust and blight in wheat [11], and Early Blight, Late Blight, and Yellow Leaf Curl Virus in tomatoes [12, 13]. Plant diseases significantly threaten agricultural productivity by causing substantial annual food losses and reducing product quality [14].

The development of deep learning, particularly transfer learning, has opened up significant opportunities in image-based plant disease detection [15]. Models such as ResNet, VGG, and MobileNet have been shown to achieve high accuracy in leaf disease classification [16, 17]. However, most research still focuses on a single crop species or relatively large and clean datasets. Real challenges such as data limitations [18], symptom variation between species, and the need for implementation on resource-limited devices remain largely unexplored [19].

Based on the identified research gaps, this study proposes an efficient and lightweight transfer learning based hybrid deep learning framework with approximately 10 million parameters. The proposed model integrates DenseNet121 and a scaled MobileNetV2 ($\alpha = 0.35$), enhanced by attention mechanisms, batch normalization, Swish activation, and Grad-CAM++ for improved interpretability. Experiments are conducted on three publicly available datasets, cassava, wheat, and tomato, to evaluate cross-crop generalization capability. The main contributions of this study are summarized as follows:

1. a parameter-efficient hybrid architecture (~10 million parameters) that balances representational capacity and computational efficiency by combining DenseNet121 and MobileNetV2;

2. a structured feature fusion strategy employing 1×1 convolutional compression to reduce feature redundancy while preserving discriminative information;
3. the integration of attention mechanisms, namely Attentive Transition and Spatial Attention, to enhance disease symptom localization;
4. a multi-crop evaluation protocol across cassava, wheat, and tomato datasets, addressing the limitation of single-crop focus in prior studies; and
5. edge-device deployment validation using TensorFlow Lite (TFLite), including latency and FPS analysis, which is rarely considered in existing works.

Thus, this study emphasizes not only accuracy but also efficiency and scalability, making it more relevant for practical applications in artificial intelligence-based plant disease detection systems.

2. RELATED WORKS

Convolutional Neural Networks (CNN) have been used extensively to identify cassava diseases, including Cassava Brown Streak Disease (CBSD) and Cassava Mosaic Disease (CMD). Oyewola and associates. Oyewola et al. [20] employed a residual CNN and reported 96.75% accuracy with an F1-score of 0.97. Ahishakiye et al. [21] applied ensemble model based on Learning Vector Quantization Algorithms (LVQ), achieving accuracy 82%. Abayomi-Alli et al. obtained 99.70% accuracy using modified MobileNetV2 neural network [22]. Despite these promising results, most cassava studies rely on relatively small datasets and seldom consider model efficiency on resource-limited devices.

Li et al. [23] compared VGG16, Inception, ResNet50, MobileNet, and DenseNet for wheat leaf disease detection and reported strong classification performance with 98.60%. Ashraf et al. [24] achieved 93% accuracy using CNNs with resampling techniques. Sharma and Sethi [3] reviewed deep CNN approaches and reported an F1-score of 0.88. In another study, Ju et al. used multispectral Unmanned Aerial Vehicle (UAV) imagery with a Back Propagation Neural Network (BPNN) to identify rust infections, achieving a positive predictive value of 92.20% [25]. These findings confirm CNNs' effectiveness for wheat disease recognition, though performance often depends heavily on dataset quality.

Tomato disease detection has become a focus of numerous studies due to the high economic value of this commodity. Bouni et al. [26] successfully classified ten tomato diseases with 99.90% accuracy using ResNet with RmsProp optimizer. Sowmiya et al. [27] trained a CNN with 800 images and achieved high performance with 97.03% accuracy. Ahmed et al. [28] utilized the lightweight MobileNet architecture and reported 99.30% accuracy. Transfer learning using eleven CNN models also achieved 98.40% accuracy [29], while other studies using DenseNet reported up to 99.30% [30]. Modified Resnet50 demonstrated results exceeding 99.49% [31], demonstrating its potential for real-time applications.

Table 1. Comparative analysis of recent deep learning approaches for plant leaf disease recognition

Author	Crop	Architecture / Method	Dataset (Size)	Main Findings
[20]	Cassava	DRNN	Cassava dataset (5,656 images)	96.75% (Accuracy)
[21]	Cassava	Deep Gaussian TL	Cassava dataset (size not specified)	82% (Accuracy)
[22]	Cassava	Modified MobileNetV2	Cassaca dataset (94,350 images)	99,70% (Accuracy)
[23]	Wheat	DQN algorithm	Wheat dataset (12,000 images)	98.60% (Accuracy)
[24]	Wheat	CNN + Resampling	Wheat dataset (450 images)	93% (Accuracy)
[25]	Wheat	BPNN Method	Wheat dataset (size not specified)	92.20% (Accuracy)
[26]	Tomato	ResNet with RmsProp optimizer	Tomato dataset (7,301 images)	99.90% (Accuracy)
[27]	Paddy	InceptionNet3	Paddy dataset (18,800 images)	97.03% (Accuracy)
[28]	Tomato	MobileNetV2	Tomato dataset (18,160 images)	99.30% (Accuracy)
[29]	Tomato	TL using 7 CNN models	Tomato dataset (18,160 images)	99.40% (Accuracy)
[30]	Tomato	DenseNet201	Tomato dataset (22,930 images)	99.30% (Accuracy)
[31]	Tomato, Rice, Cassava, Orange, Peach, Potato	Modified Resnet50	Multi-crop dataset (43,869 images)	99.49% (Accuracy)
[32]	Pepper bell, Potato, Tomato	ViT (TLMViT)	Multi-crop dataset (20,638 images)	98.81% (Accuracy)
[33]	Wheat	ResNet50	Wheat dataset (20,421 images)	96.30% (Accuracy)
[34]	Tomato	DenseNet	Tomato dataset (3,000 images)	91.33% (Accuracy)
[36]	Wheat, Cassava	LeafDoc-Net	Multi-crop dataset (635 images)	99.41% (Accuracy)
[37]	Tomato	HCA and CSA	Tomato dataset (19,969 images)	98.71% (Accuracy)
[38]	Tomato	MobileNet2	Tomato dataset (14,529 images)	99.01% (Accuracy)

Several studies have attempted to combine the strengths of different architectures. Tabbakh and Barpanda [32] introduced Transfer Learning Model with Vision Transformer (TLMViT) with 98.81% accuracy. Mandava et al. [33] reported EfficientNetB3 with 96.30% accuracy on wheat diseases. Several studies have also explored hybrid models based on DenseNet and MobileNet, achieving 91–99% accuracy while accelerating training [34-36]. However, most studies have not focused on computational efficiency and deployment on edge devices.

The efficacy of deep learning and transfer learning in the classification of plant leaf diseases across different commodities has been shown in a number of earlier studies. However, most studies still focus on a single crop type or dataset, with little attention paid to computational efficiency and readiness for implementation on edge devices. A summary of related studies and their key results is presented in Table 1. As shown in Table 1, CNN based and transfer learning models generally achieve strong classification performance, with reported accuracies often exceeding 90%. ResNet and DenseNet have proven superior in accuracy, while MobileNet and its variants excel in speed and computational efficiency. Several recent approaches, such as Vision Transformer (ViT) and EfficientNet, have also been reported to achieve high performance on certain datasets. However, several limitations remain apparent:

1. Limited focus on a single crop commodity. Most studies only test models on a single dataset (e.g., tomato or wheat), so their generalizability across crop species remains largely unexplored.
2. Lack of attention to model efficiency. Many studies produce models with a large number of parameters, which are less suitable for implementation in resource-constrained agricultural systems.
3. Lack of interpretability. Only a few studies emphasize visual explanations (e.g., Grad-CAM++) to increase practitioners' confidence in the classification results.

Based on this gap, this study proposes an efficient (~10 million parameters) hybrid DenseNet121 and MobileNetV2. This model is tested not only on one commodity, but on three different datasets (cassava, wheat, and tomato), while integrating Grad-CAM++ to support interpretability of the results. Additionally, an evaluation of edge deployment using TFLite is conducted to verify that it is ready for real-time implementation on devices with limited resources.

3. PROPOSED METHOD

This study aims to develop an efficient transfer learning-based deep learning model for multiclass classification of leaf diseases in three main commodities: cassava, wheat, and tomato. The architecture of the proposed hybrid DenseNet121 and MobileNetV2 framework, shown in Figure 1, is as follows: input images are processed in parallel by two backbones, enhanced with attention mechanisms, compressed using 1 × 1 convolutions, aggregated via Global Average Pooling (GAP), and fused for classification.

3.1 Dataset

The datasets used consisted of three public sources: the Cassava Leaf Disease Dataset [39], which includes Cassava Blight (CB), Cassava Mosaic (CM), and Cassava Healthy (CH). The Wheat Leaf Disease Dataset [40], which includes Wheat Septoria (WS), Wheat Stripe Rust (WSR), and Wheat Healthy (WH). The dataset for Tomato leaf disease [41], which includes Tomato Bacterial Spot (TBS), Tomato Early Blight (TEB), Tomato Late Blight (TLB), Tomato Leaf Mold (TLM), Tomato Mosaic Virus (TMV), Tomato Septoria Leaf Spot (TSLS), Tomato Spider Mites Two-spotted Spider Mite (TSMTSM), Tomato Target Spot (TTS), Tomato Yellow Leaf Curl Virus (TYLCV), and Tomato Healthy (TH). Thirteen disease classes and one healthy class comprised the total of 10,635 images. To ensure the preservation of the original class distribution, the dataset was divided into training (80%) and testing (20%) sets using a stratified sampling approach. This strategy was applied consistently across all datasets and is particularly important for handling class imbalance, such as the cassava dataset with relatively few samples per class and the tomato dataset with substantially larger class sizes. Example images are shown in Figure 2. The detailed distribution of image counts is presented in Table 2.

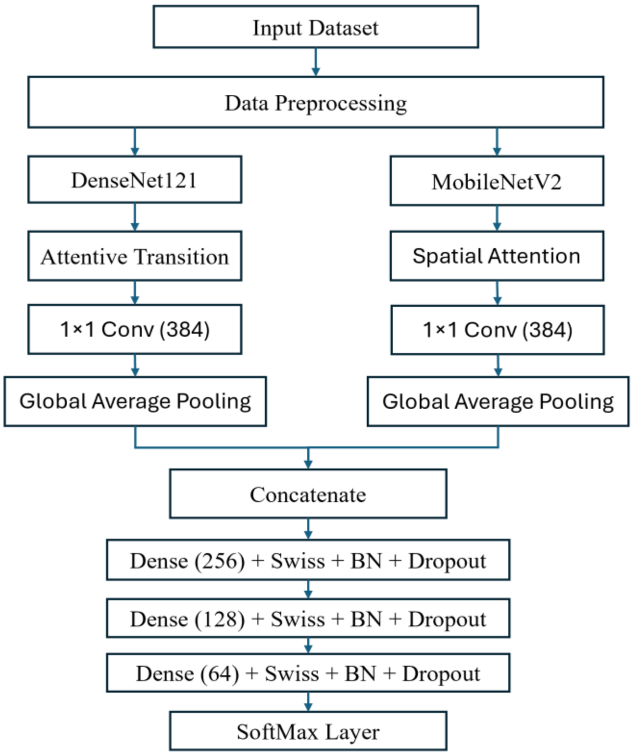


Figure 1. Design of the proposed hybrid deep learning framework combining DenseNet121 and MobileNetV2 with attention-based feature fusion

3.2 Data preparation and augmentation pipeline

All input images were standardized to a 224 × 224 resolution [42] and scaled to the [0, 1] range via division by 255. An augmentation strategy was subsequently employed to suppress overfitting [43], applying transformations such as rotation (±30°), spatial shifts (0.2), shear and zoom (0.2), horizontal flipping, and brightness variation (0.7–1.3). After augmentation, the amount of training data increased significantly, increasing diversity.

3.3 Hybrid model architecture

The proposed hybrid model integrates two complementary

CNN backbones: DenseNet121 and MobileNetV2 (α = 0.35), aiming to balance representational richness and computational efficiency. DenseNet121 is employed to capture detailed and hierarchical feature representations through dense connectivity, while MobileNetV2 provides lightweight and efficient feature extraction using depthwise separable convolutions.

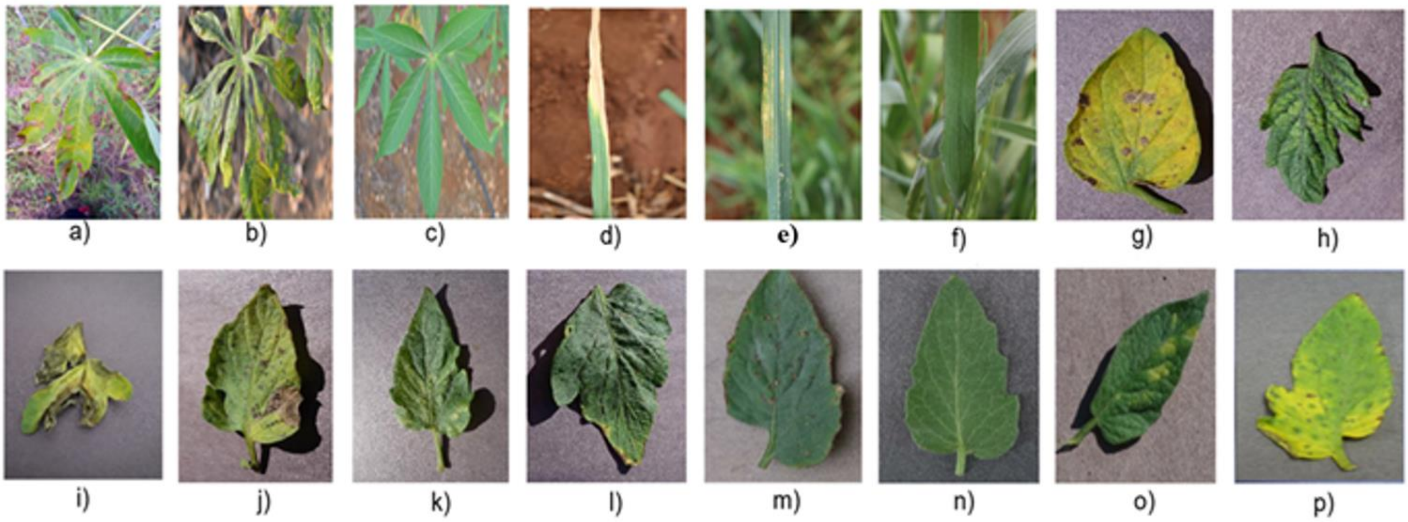


Figure 2. a) CB, b) CM, c) CH, d) WS, e) WSR, f) WH, g) TSLS, h) TMV, i) TEB, j) TSMTSM, k) TTS, l) TBS, m) TLM, n) TH, o) TLM, p) TYLCV

Table 2. Multi-crop dataset collection: number of training and testing images per leaf disease class [39-41]

Dataset	Class	Training	Testing
Cassava	CB	39	10
	CM	70	18
	CH	73	18
Wheat	WH	81	21
	WS	78	19
	WSR	169	39
Tomato	TBS	800	200
	TEB	800	200
	TLB	800	200
	TLM	800	200
	TMV	800	200
	TSLS	800	200
	TSMTSM	800	200
	TTS	800	200
	TYLCV	800	200
	TH	800	200

Feature extraction is performed at the final convolutional stage of each backbone. For DenseNet121, features are extracted after the last dense block (conv5_block16_concat), followed by a newly defined Attentive Transition module that applies channel-wise attention to emphasize disease-relevant features. For MobileNetV2 (α = 0.35), features are taken from the output of the final convolutional layer (Conv_1) and enhanced using a Spatial Attention module implemented via a 1 × 1 convolution followed by a sigmoid activation.

In both branches, the resulting feature maps are compressed to 384 channels using 1 × 1 convolutions and then individually aggregated using GAP. The pooled feature vectors from DenseNet121 and MobileNetV2 are subsequently concatenated to form a unified representation. The classification head consists of fully connected layers with dimensions 256 → 128 → 64, where the Swish activation function is consistently applied in all layers.

1. Swish Activation Function

$$f(x) = x.\sigma(x) = x.\frac{1}{1+e^{-x}} \tag{1}$$

where, x: input neuron, $\sigma(x)$: sigmoid function.

Swish is a non-monotonic activation function that is smoother than ReLU, thus improving training stability.

Batch Normalization and Dropout (rate = 0.5) are employed after each dense layer to improve training stability and reduce overfitting. Finally, a Softmax layer is used to produce class probabilities according to the number of disease categories.

2. Softmax

$$Softmax\ (z_i) = \frac{e^{zi}}{\Sigma_j e^{zj}} \tag{2}$$

where, z: logit (output value before normalization) for class c, K: number of classes, $P(y = c|x)$: probability that input x belongs to class c. Softmax ensures that the output is a probability distribution with a sum of 1.

Attentive Transition aims to improve feature retention in the transition layers of DenseNet architectures. In traditional DenseNets, these layers reduce spatial resolutions and channel dimensions, which can inadvertently lead to the loss of critical diagnostic information. By incorporating channel and spatial attention modules within these transition layers, the model dynamically assesses and retains important features. This is achieved by calculating channel importance scores through GAP and leveraging gating mechanisms, alongside generating spatial saliency maps via convolutional aggregations of feature responses. This dual mechanism creates a sophisticated attention mask that modulates feature maps prior to pooling, ultimately enhancing the model's capability to discern and retain significant pathognomonic features, thereby improving generalization across dense blocks [44-46].

Furthermore, Spatial Attention complements the Attentive

Transition by specifically targeting symptom-localizing areas within input images. This mechanism develops 2D attention maps that amplify features related to diagnostic symptoms such as lesions or necrosis while reducing the influence of irrelevant background noise. The creation of these attention maps through lightweight convolutional blocks leads to element-wise multiplication with the feature tensor, allowing the model to focus more sharply on clinically significant areas [46]. Research indicates that employing spatial attention enhances model accuracy and facilitates interpretability in the predictions made by neural networks [47, 48].

Moreover, the advantages of incorporating attention mechanisms extend beyond mere accuracies; they contribute to a better understanding of model behaviors. As shown in various studies, models utilizing attention can derive more interpretable insights, highlighting which areas of the input data were crucial for classification. This is particularly valuable in plant pathology, as determining symptoms associated with diseases can inform effective mitigation strategies and agricultural practices [49, 50].

Synthesizing Attentive Transition and Spatial Attention into the architectures of DenseNet121 and MobileNetV2 provides a robust framework for diagnosing plant diseases. This hybrid model significantly enhances the ability to extract and retain relevant features while offering clearer insights into the disease classification process. Implementing these advanced attention mechanisms aligns with the current trends in deep learning research that emphasize accuracy and interpretability [51-54].

3.4 Training process

Model training was conducted using a TensorFlow/Keras-based computing environment with Graphics Processing Unit (GPU) acceleration via CUDA on an NVIDIA T4 device with 15 GB of memory, enabling efficient and rapid training. Adam was used as the optimizer with an initial learning rate of 1×10^{-5}, chosen for its ability to adaptively adjust the learning rate during training, thus promoting stable convergence. The batch size was set at 8, aligning with GPU memory limitations and ensuring good gradient quality during weight updates. The training process ran for a maximum of 50 epochs, but included an early stopping mechanism that monitored the validation loss. This meant that training would be stopped early if there was no decrease in the validation loss over several consecutive epochs, preventing overfitting and saving computational time. To improve model generalization, a dropout regularization technique with a dropout rate of 0.5 was applied, which randomly deactivates half of the neurons in a given layer during training, forcing the network to learn a more robust representation. For performance comparison, two specialized models for leaf disease detection were used: LeafDoc-Net and its modified variant, LeafDoc-Net ReLU. This comparison aims to comprehensively evaluate the performance of the proposed models, including accuracy, efficiency, and generalization capabilities for leaf image data.

3.5 Performance metrics

Model evaluation was performed using standard multiclass classification metrics: accuracy, precision, recall, F1-score, and Area Under the Curve (AUC). Analysis was performed using a confusion matrix to examine the distribution of predictions between classes. Some of the mathematical formulations used are as follows:

3. Accuracy

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{3}$$

with: TP: true positive, TN: true negative, and FP: false positive, FN: false negative.

4. Precision

$$Precision = \frac{TP}{TP + FP} \tag{4}$$

with: TP: true positive (correct prediction for the positive class), FP: false positive (incorrect prediction, thought to be positive when it is actually negative).

Precision measures the extent to which the model accurately predicts the positive class.

5. Recall

$$Recall = \frac{TP}{TP + FN} \tag{5}$$

with: TP: true positive, FN: false negative (incorrect prediction, mistaken for a negative class when it is actually positive). Recall measures the extent to which the model is able to find all positive examples.

6. F1-Score

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{6}$$

$F1$ is the harmonic mean of precision and recall, useful when the class distribution is imbalanced.

3.6 Efficiency and deployment testing

In addition to accuracy, this study also measures model efficiency through the number of parameters and inference latency. The model was converted to TFLite for testing on edge devices. The parameters evaluated included average latency (ms), p95 latency, p99 latency, and frames per second (FPS) to assess readiness for real-time implementation.

4. RESULT AND ANALYSIS

4.1 Model performance comparison

The proposed hybrid model was evaluated on three publicly available datasets cassava, wheat, and tomato, and compared against two closely related baseline architectures, namely LeafDoc-Net and LeafDoc-Net ReLU. The quantitative performance metrics, including Accuracy, Precision, Recall, AUC, and the number of model parameters, are summarized in Tables 3-5. On the cassava dataset (Table 3), the proposed model achieves perfect classification performance, with Accuracy, Precision, Recall, and AUC all equal to 1.0. While this performance matches that of LeafDoc-Net, the proposed architecture requires fewer parameters (approximately 10 million compared to 12 million), demonstrating improved parameter efficiency without sacrificing predictive accuracy. On the wheat dataset (Table 4), LeafDoc-Net attains the highest accuracy (0.975), whereas the proposed model achieves a slightly lower accuracy (0.950). Despite this marginal difference, the proposed model maintains competitive performance while using a more compact architecture, highlighting its robustness on a dataset characterized by higher classification difficulty.

Table 3. Comparative evaluation of pre-trained and proposed models for cassava leaf disease dataset

Model	Accuracy	Precision	Recall	Area Under the Curve (AUC)	Number of Parameters (Million)
LeafDoc-Net_Relu [36]	0.9545	0.9545	0.9530	0.9964	12
LeafDoc-Net [36]	1	1	1	1	12
Proposed Method	**1**	**1**	**1**	**1**	**10**

Table 4. Comparative evaluation of pre-trained and proposed models for wheat leaf disease dataset

Model	Accuracy	Precision	Recall	Area Under the Curve (AUC)	Number of Parameters (Million)
LeafDoc-Net_Relu [36]	0.9500	0.9601	0.9401	0.9800	12
LeafDoc-Net [36]	0.9750	0.9743	0.9747	0.9932	12
Proposed Method	**0.9500**	**0.9512**	**0.9500**	**0.9951**	**10**

Table 5. Comparative evaluation of pre-trained and proposed models for tomato leaf disease dataset

Model	Accuracy	Precision	Recall	Area Under the Curve (AUC)	Number of Parameters (Million)
LeafDoc-Net_Relu [36]	0.953	0.9733	0.9530	0.9897	12
LeafDoc-Net [36]	0.9941	0.9940	0.9940	0.9999	12
Proposed Method	**0.9960**	**0.9960**	**0.9960**	**0.9999**	**10**

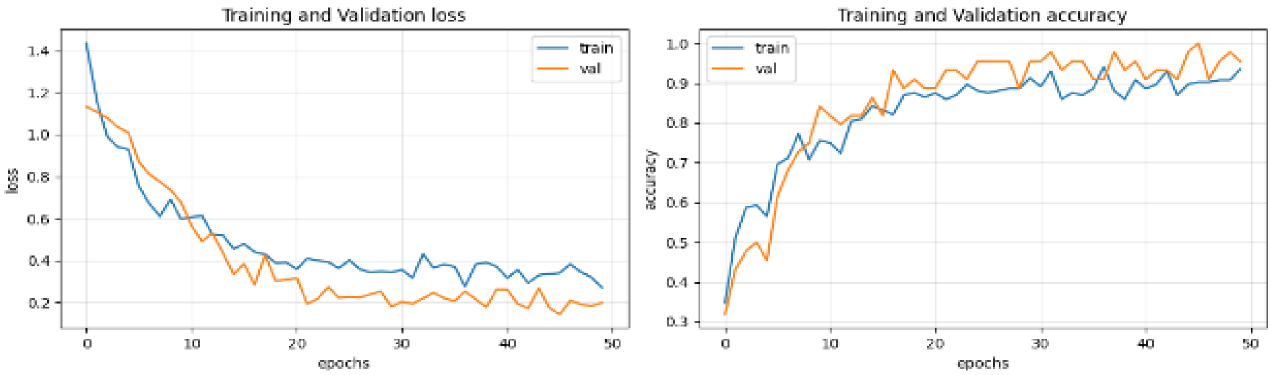
On the tomato dataset (Table 5), which involves a more complex multi-class classification task with ten disease categories, the proposed model achieves the highest accuracy (0.996), outperforming both LeafDoc-Net (0.9941) and LeafDoc-Net ReLU (0.953). This result indicates that the DenseNet–MobileNet hybrid architecture is particularly effective in scenarios with increased inter-class variability and complex symptom patterns. Overall, these results demonstrate that the proposed model consistently achieves a favorable balance between classification performance and model compactness across all evaluated datasets.

4.2 Training curve analysis

Figures 3-5 present the training and validation curves of the proposed model for the cassava, wheat, and tomato datasets. Across all experiments, the model exhibits rapid and stable convergence, typically within the first 50 epochs, with a small and consistent gap between training and validation curves. This behavior indicates that the applied regularization strategies, including dropout and data augmentation, effectively mitigate overfitting. The AUC values approach 1.0 at early stages of training, reflecting strong inter-class separability and robust feature representation. Notably, for the tomato dataset, the validation performance occasionally exceeds the training performance, suggesting strong generalization capability despite the increased complexity associated with a larger number of disease classes.

Further insights are provided by the confusion matrices shown in Figures 6(a)-(b) and 7. For the cassava dataset, the proposed model achieves perfect classification across all classes, with no observed misclassifications. This highlights the effectiveness of the hybrid architecture and attention mechanisms in capturing distinct disease characteristics. On the wheat dataset, the model performs very well for the healthy class; however, minor misclassifications occur between the Septoria and Stripe Rust classes. These errors are primarily attributed to the high visual similarity between the two diseases, both of which exhibit elongated lesion patterns with comparable color and texture features. Additionally, the wheat dataset shows class imbalance, where the Stripe Rust class contains substantially more training samples than the Septoria class, potentially biasing the learned decision boundaries.

For the tomato dataset, the classification task is inherently more challenging due to the presence of ten distinct disease classes. Despite this complexity, the proposed model maintains high accuracy and AUC values, demonstrating its robustness and ability to generalize across diverse disease categories.



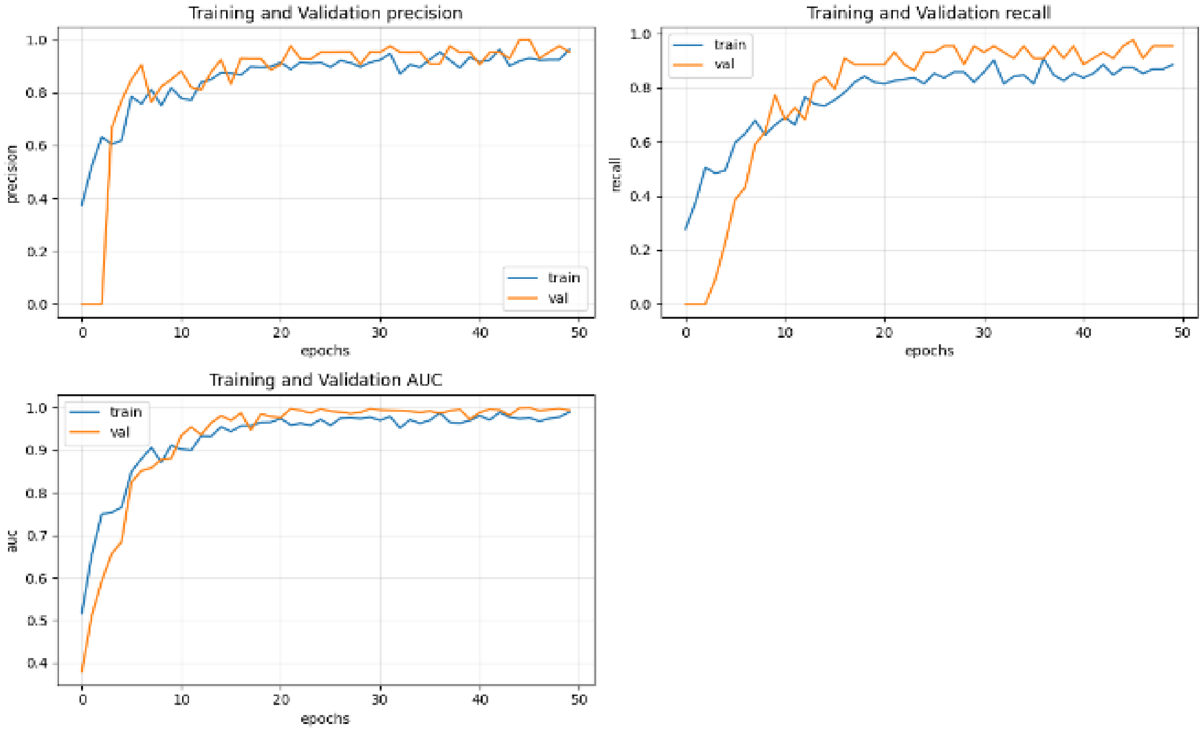


Figure 3. Training dynamics and validation performance trends of the proposed model, illustrating loss, accuracy, precision, recall, and Area Under the Curve (AUC) across epochs on the cassava leaf disease dataset

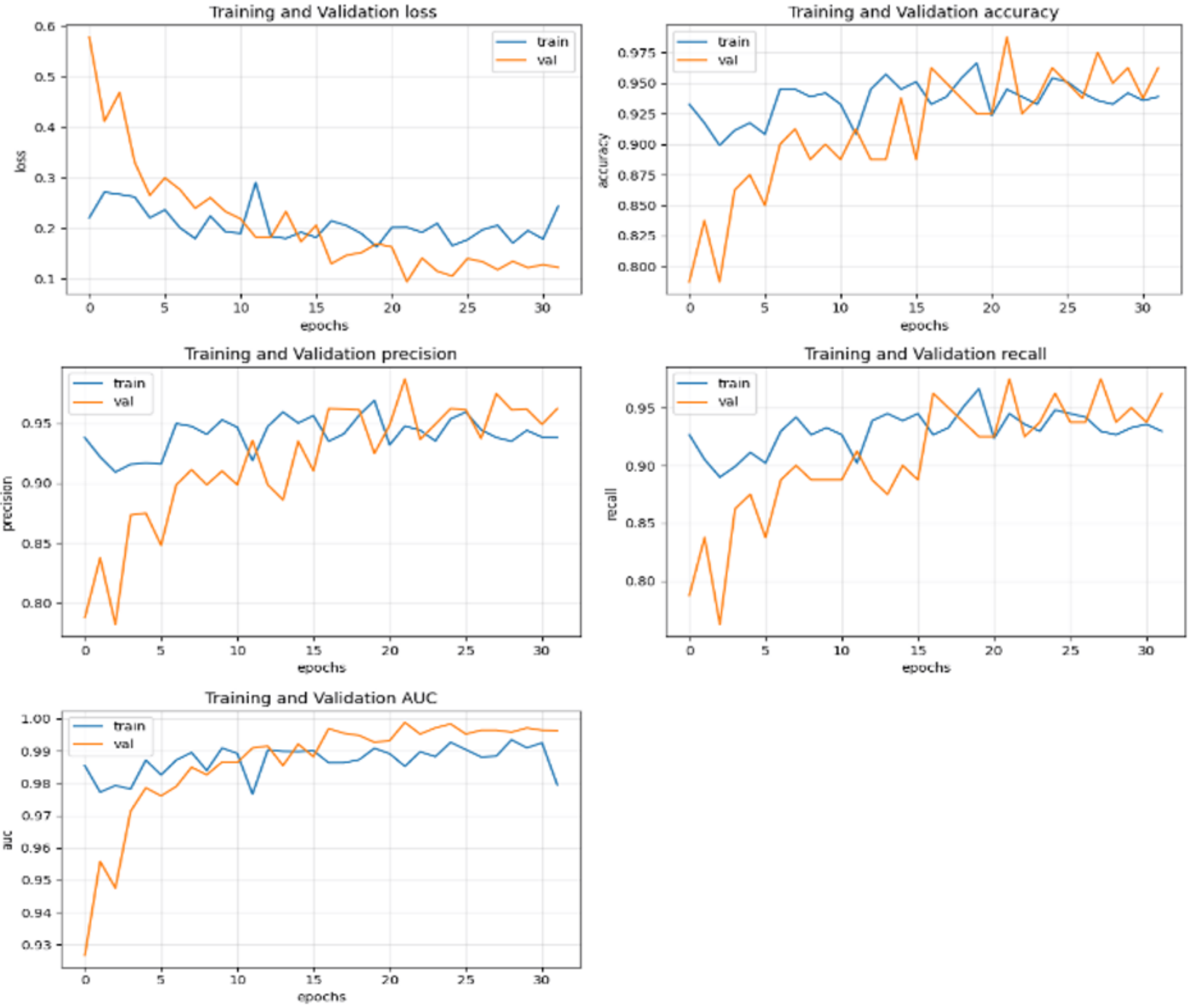


Figure 4. Training dynamics and validation performance trends of the proposed model, illustrating loss, accuracy, precision, recall, and Area Under the Curve (AUC) across epochs on the wheat leaf disease dataset

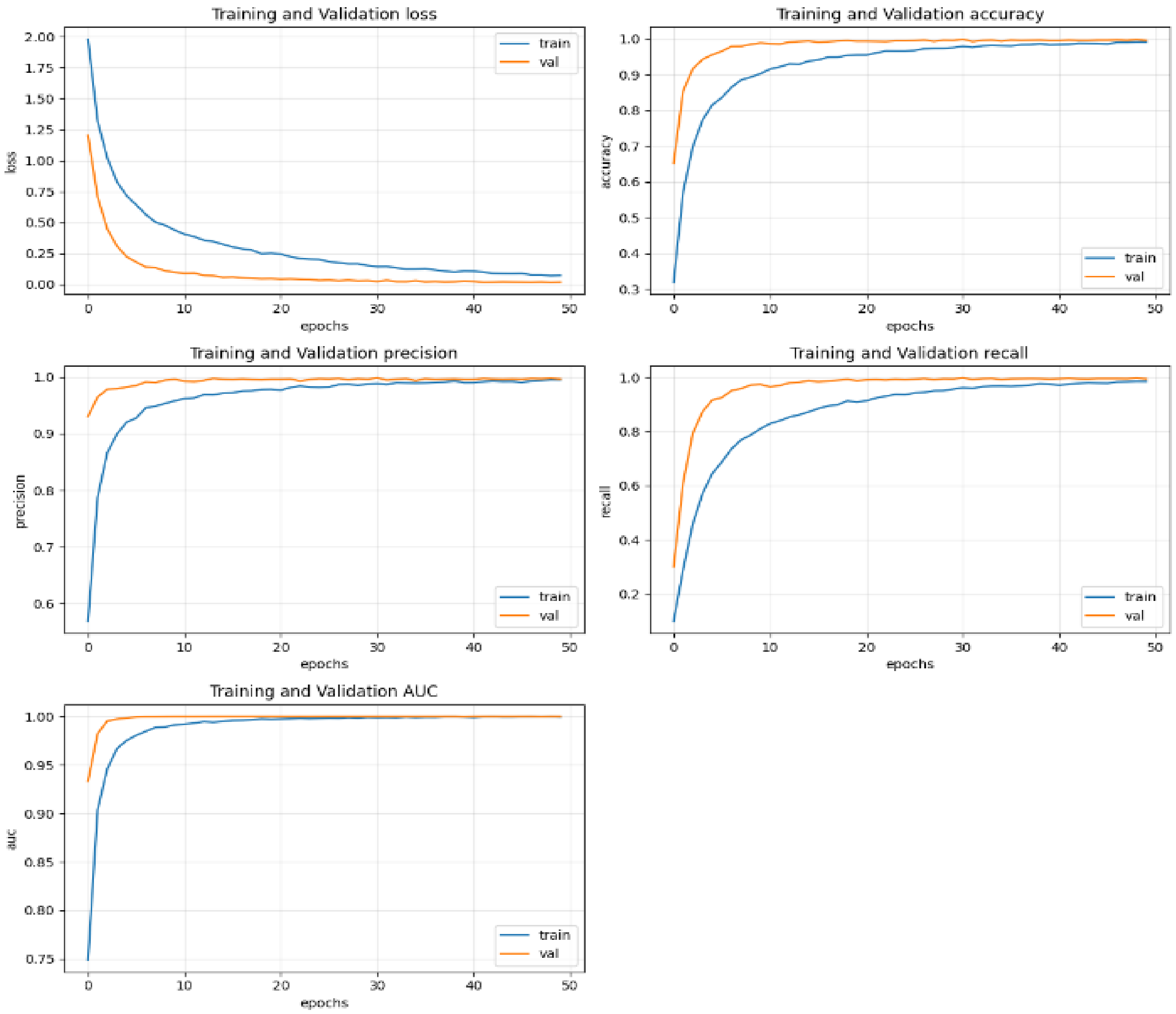
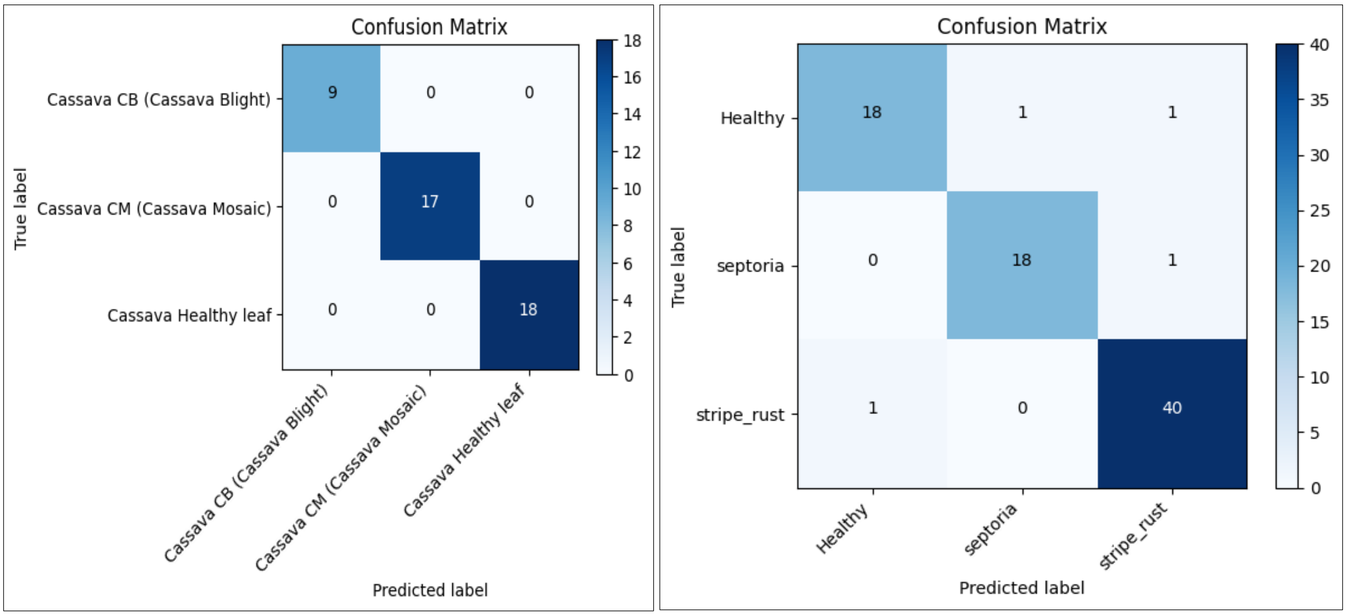
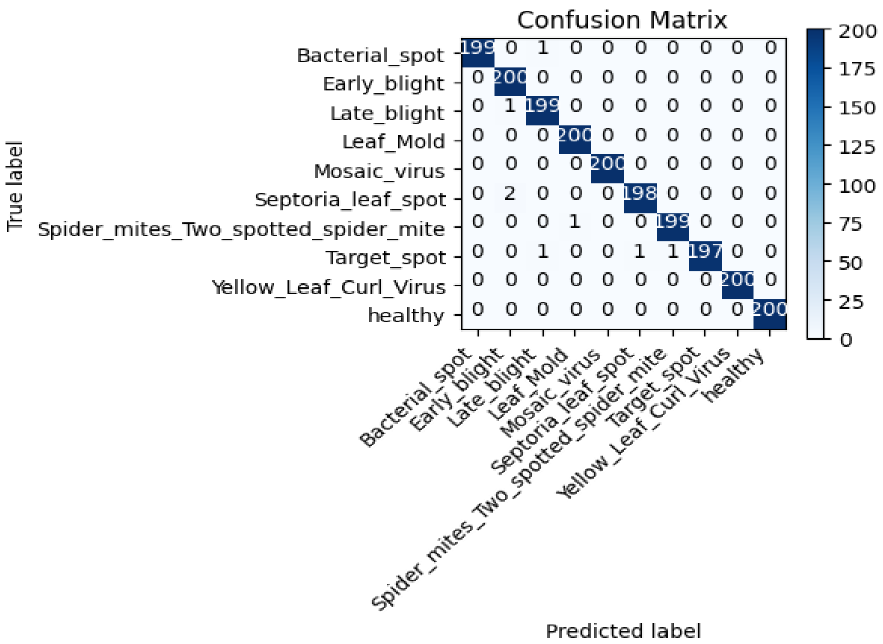


Figure 5. Training dynamics and validation performance trends of the proposed model, illustrating loss, accuracy, precision, recall, and Area Under the Curve (AUC) across epochs on the tomato leaf disease dataset



(a) (b)

Figure 6. Confusion matrix of LeafDoc-Net & LeafDoc-Net ReLU architectures on leaf disease datasets a). Cassava b) Wheat



	Bacterial_spot	Early_blight	Late_blight	Leaf_Mold	Mosaic_virus	Septoria_leaf_spot	Spider_mites_Two_spotted_spider_mite	Target_spot	Yellow_Leaf_Curl_Virus	healthy
Bacterial_spot	199	0	1	0	0	0	0	0	0	0
Early_blight	0	200	0	0	0	0	0	0	0	0
Late_blight	0	1	199	0	0	0	0	0	0	0
Leaf_Mold	0	0	0	200	0	0	0	0	0	0
Mosaic_virus	0	0	0	0	200	0	0	0	0	0
Septoria_leaf_spot	0	2	0	0	0	198	0	0	0	0
Spider_mites_Two_spotted_spider_mite	0	0	0	1	0	0	199	0	0	0
Target_spot	0	0	1	0	0	1	1	197	0	0
Yellow_Leaf_Curl_Virus	0	0	0	0	0	0	0	0	200	0
healthy	0	0	0	0	0	0	0	0	0	200

Figure 7. Confusion matrix of the proposed architecture on the tomato leaf disease dataset

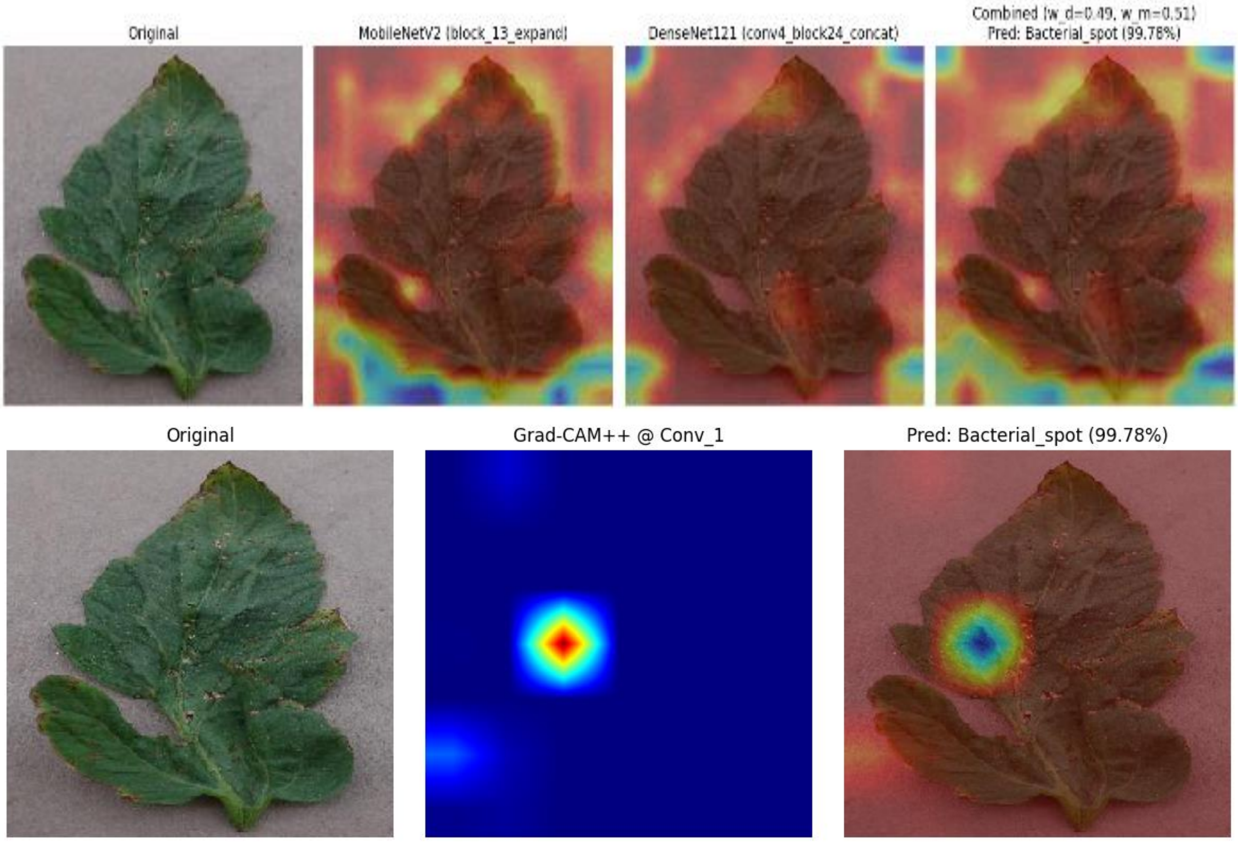


Figure 8. Comparison between the original image and the results of the Grad-CAM++ interpretation

4.3 Interpretation with Grad-CAM++

To improve model interpretability, Grad-CAM++ visualizations were generated for representative test samples, as shown in Figure 8. The resulting heatmaps indicate that the proposed model consistently focuses on disease-relevant regions of the leaf, rather than background elements, supporting the biological plausibility of the learned representations. For example, in wheat leaf samples affected by Stripe Rust, the highlighted regions correspond to the characteristic yellow–orange linear lesions associated with the disease. Similar attention patterns are observed across cassava and tomato datasets, where the model emphasizes lesion boundaries, discoloration, and texture irregularities that are diagnostically meaningful. These visual explanations provide additional confidence that the model's predictions are driven by relevant plant pathology features, which is particularly important for increasing user trust and facilitating adoption in real-world agricultural applications.

4.4 Efficiency and deployment testing

Table 6 shows the edge deployment performance using TFLite. The proposed model has an average latency of 46–55 ms and a throughput of >20 FPS, which meets the requirements for real-time applications. These results are important because many IoT-based smart agriculture systems have limited computing power.

Overall, all three models performed very well on average for deployment on edge devices. Data B (tomato) and Data C (cassava) were the best performers. These two models performed nearly identically and were significantly more optimal for deployment on edge devices compared to Data A (wheat). Consistency is a Plus: All three models were very consistent (good p95 values). Data B and C excelled in their higher average speed and better worst-case (p99) performance. Edge-Ready Model: Data D (Cassava) confirmed the positive trend of Data C. With an FPS of ~21 and consistent latency, this model is well-suited for real-time applications on resource-constrained devices.

Table 6. Comparison of TFLite (CPU) performance for edge readiness

Metric	Wheat (Data A)	Tomato (Data B)	Cassava (Data C)
Avg Latency	54.67 ms	46.88 ms	47.74 ms
p95 Latency	55.92 ms	54.10 ms	55.06 ms
p99 Latency	85.07 ms	75.24 ms	74.14 ms
FPS	18.3	21.3	20.9

4.5 Critical discussion

This study demonstrates that the proposed hybrid DenseNet121-MobileNetV2 framework achieves a favorable balance between accuracy, efficiency, and deployment readiness. With approximately 10 million parameters, the model is notably lighter than LeafDoc-Net (12 million parameters), while maintaining competitive performance across all evaluated datasets and achieving superior results on the tomato dataset. A key strength of the proposed approach lies in its generalizability across multiple crop species. Unlike most existing studies that focus on a single crop, the proposed model was evaluated on three different commodities, cassava, wheat, and tomato and demonstrated stable performance across these domains.

From a practical perspective, the TFLite deployment results confirm that the model is suitable for edge-device implementation, achieving low inference latency and high FPS. This characteristic is particularly important for real-world precision agriculture applications, where computational resources are often limited.

Nevertheless, several limitations should be acknowledged. The high performance achieved on public benchmark datasets may be influenced by relatively clean imaging conditions. Future work should therefore evaluate the model on in-the-wild datasets that include variations in lighting, background clutter, and image quality. Additionally, the observed confusion between Septoria and Stripe Rust in the wheat dataset indicates that further improvements are possible. Future research may explore refined attention mechanisms or transformer-based architectures to better capture subtle inter-class differences and mitigate the effects of class imbalance.

4.6 Research implications

These results demonstrate that the hybrid architecture not only improves accuracy but also balances efficiency, interpretability, and deployment readiness. This makes the proposed approach more relevant for practical implementation in AI-based plant disease detection systems, especially in developing countries with limited resources. In particular, the emphasis on lightweight architectures with verified low-latency inference makes the proposed approach especially relevant for resource-constrained agricultural environments, such as those commonly encountered in developing countries.

5. CONCLUSION

This study proposes a transfer learning-based hybrid architecture combining DenseNet121 and MobileNetV2 for multi-class classification of plant leaf diseases. With a relatively small number of parameters (~10 million), the model demonstrated competitive performance on three public datasets: cassava, wheat, and tomato. Experimental results recorded perfect accuracy on the cassava dataset, 95% accuracy on wheat, and a peak accuracy of 99.6% on tomato, while also being more efficient than benchmark models with larger parameters. Further evaluation using TFLite confirmed the model's readiness for implementation on edge devices with low latency (<60 ms) and high throughput (>20 FPS), enabling real-time use in precision agriculture practices. Grad-CAM++ integration also improved interpretability, ensuring model predictions are based on relevant visual features. Overall, this research offers important contributions to the development of accurate, efficient, and field-ready plant disease detection systems. Future research directions include expansion to in-the-wild datasets with diverse lighting conditions and backgrounds, integration with IoT systems or agricultural robots for automated diagnosis in the field, and exploration of novel transformer- or attention-refinement-based architectures to improve accuracy on difficult-to-distinguish disease classes. Thus, this research makes a significant contribution to the field of precision agriculture through the development of models that are not only accurate, but also efficient, well-interpretable, and ready for use in real applications.

REFERENCES

[1] Sundararaman, B., Jagdev, S., Khatri, N. (2023). Transformative role of artificial intelligence in advancing sustainable tomato (Solanum lycopersicum) disease management for global food security: A comprehensive review. Sustainability, 15(15): 11681. https://doi.org/10.3390/su151511681

[2] Shigaki, T. (2016). Cassava: The nature and uses. In Encyclopedia of Food and Health, pp. 687-693. https://doi.org/10.1016/B978-0-12-384947-2.00124-0

[3] Sharma, T., Sethi, G.K. (2024). Improving wheat leaf disease image classification with point rend segmentation technique. SN Computer Science, 5: 244. https://doi.org/10.1007/s42979-023-02571-w

[4] Melanson, R.A. Common diseases of tomatoes. https://extension.msstate.edu/publications/common-

diseases-tomatoes, accessed on Aug. 19, 2025.

[5] Amelework, A.B., Bairu, M.W., Maema, O., Venter, S.L., Laing, M. (2021). Adoption and promotion of resilient crops for climate risk mitigation and import substitution: A case analysis of cassava for South African agriculture. Frontiers in Sustainable Food Systems, 5: 617783. https://doi.org/10.3389/fsufs.2021.617783

[6] Nandudu, L., Sheat, S., Winter, S., Ogbonna, A., Kawuki, R., Jannink, J.L. (2024). Genetic complexity of cassava brown streak disease: Insights from qPCR-based viral titer analysis and genome-wide association studies. Frontiers in Plant Science, 15: 1365132. https://doi.org/10.3389/fpls.2024.1365132

[7] Singh, V., Sharma, N., Singh, S. (2020). A review of imaging techniques for plant disease detection. Artificial Intelligence in Agriculture, 4: 229-242. https://doi.org/10.1016/j.aiia.2020.10.002

[8] Oliveira, M., Teixeira, A., Barreto, G., Lima, C. (2024). GamaNNet: A novel plant pathologist-level CNN architecture for intelligent diagnosis. AgriEngineering, 6(3): 2623-2639. https://doi.org/10.3390/agriengineering6030153

[9] Godding, D., Stutt, R.O.J.H., Alicai, T., Abidrabo, P., Okao-Okuja, G., Gilligan, C.A. (2023). Developing a predictive model for an emerging epidemic on cassava in sub-Saharan Africa. Scientific Reports, 13: 12603. https://doi.org/10.1038/s41598-023-38819-x

[10] Ano, C.U. Ochwo-Ssemakula, M., Ibanda, A., Ozimati, A., Gibson, P., Onyeka, J., Njoku, D., Egesi, C., Kawuki, R.S. (2021). Cassava brown streak disease response and association with agronomic traits in elite Nigerian cassava cultivars. Frontiers in Plant Science, 12: 720532. https://doi.org/10.3389/fpls.2021.720532

[11] Xu, L.X. Cao, B.X., Zhao, F.J., Ning, S.Y., Xu, P., Zhang, W.B., Hou, X.G. (2023). Wheat leaf disease identification based on deep learning algorithms. Physiological and Molecular Plant Pathology, 123: 101940. https://doi.org/10.1016/j.pmpp.2022.101940

[12] Batool, A., Kim, J., Lee, S.J., Yang, J.H., Byun, Y.C. (2024). An enhanced lightweight T-Net architecture based on convolutional neural network (CNN) for tomato plant leaf disease classification. PeerJ Computer Science, 10: e2495. https://doi.org/10.7717/peerj-cs.2495

[13] Jasani, A., Dholi, M., Purkar, S. (2022). Tomato leaf disease detection. International Journal of Research in Applied Science and Engineering Technology, 10(5): 918-922. https://doi.org/10.22214/ijraset.2022.41918

[14] Gai, Y.P., Wang, H.K. (2024). Plant disease: A growing threat to global food security. Agronomy, 14(8): 1615. https://doi.org/10.3390/agronomy14081615

[15] Kabala, D.M., Hafiane, A., Bobelin, L., Canals, R. (2023). Image-based crop disease detection with federated learning. Scientific Reports, 13: 19220. https://doi.org/10.1038/s41598-023-46218-5

[16] Wagle, S.A., Harikrishnan, R., Md Ali, S.H., Faseehuddin, M. (2022). Classification of plant leaves using new compact convolutional neural network models. Plants, 11(1): 24. https://doi.org/10.3390/plants11010024

[17] Nguyen, H.T., Luong, H.H., Huynh, L.B., Hoang Le, B.Q., Doan, N.H., Dao Le, D.T. (2023). An improved mobilenet for disease detection on tomato leaves. Advances in Technology Innovation, 8(3): 192-209. https://doi.org/10.46604/aiti.2023.11568

[18] Afifi, A., Alhumam, A., Abdelwahab, A. (2020). Convolutional neural network for automatic identification of plant diseases with limited data. Plants, 10(1): 28. https://doi.org/10.3390/plants10010028

[19] Gulame, M.B., Thite, T.G., Patil, K.D. (2023). Plant disease prediction system using advance computational technique. Journal of Physics: Conference Series, 2601(1): 012031. https://doi.org/10.1088/1742-6596/2601/1/012031

[20] Oyewola, D.O., Dada, E.G., Misra, S., Damaševičius, R. (2021). Detecting cassava mosaic disease using a deep residual convolutional neural network with distinct block processing. PeerJ Computer Science, 7: e352. https://doi.org/10.7717/peerj-cs.352

[21] Ahishakiye, E., Mwangi, W., Murithi, P., Wario, R., Kanobe, F., Danison, T. (2023). An ensemble model based on learning vector quantization algorithms for early detection of cassava diseases using spectral data. In Digital-for-Development: Enabling Transformation, Inclusion and Sustainability Through ICTs. IDIA 2022. Communications in Computer and Information Science, pp. 320-328. https://doi.org/10.1007/978-3-031-28472-4_20

[22] Abayomi-Alli, O.O., Damaševičius, R., Misra, S., Maskeliūnas, R. (2021). Cassava disease recognition from low-quality images using enhanced data augmentation model and deep learning. Expert Systems, 38(7): e12746. https://doi.org/10.1111/exsy.12746

[23] Li, Y., Liu, H.B., Wei, J.L., Ma, X.M., Zheng, G., Xi, L. (2023). Research on winter wheat growth stages recognition based on mobile edge computing. Agriculture, 13(3): 534. https://doi.org/10.3390/agriculture13030534

[24] Ashraf, M., Abrar, M., Qadeer, N., Alshdadi, A.A., Sabbah, T., Khan, M.A. (2023). A convolutional neural network model for wheat crop disease prediction. Computers, Materials & Continua, 75(2): 3867-3882. https://doi.org/10.32604/cmc.2023.035498

[25] Ju, C.X., Chen, C., Li, R., Zhao, Y.Y., Zhong, X.C., Sun, R.L., Liu, T., Sun, C.M. (2023). Remote sensing monitoring of wheat leaf rust based on UAV multispectral imagery and the BPNN method. Food and Energy Security, 12(4): e477. https://doi.org/10.1002/fes3.477

[26] Bouni, M., Hssina, B., Douzi, K., Douzi, S. (2023). Impact of pretrained deep neural networks for tomato leaf disease prediction. Journal of Electrical and Computer Engineering, 2023(1): 5051005. https://doi.org/10.1155/2023/5051005

[27] Sowmiya, B., Saminathan, K., Devi, M.C. (2024). An ensemble of transfer learning based InceptionV3 and VGG16 models for paddy leaf disease classification. ECTI Transactions on Computer and Information Technology, 18(1): 89-100.

[28] Ahmed, S., Hasan, M.B., Ahmed, T., Karim Sony, M.R., Kabir, M.H. (2022). Less is more: Lighter and faster deep neural architecture for tomato leaf disease classification. IEEE Access, 10: 68868-68884. https://doi.org/10.1109/access.2022.3187203

[29] Khasawneh, N., Faouri, E., Fraiwan, M. (2022). Automatic detection of tomato diseases using deep transfer learning. Applied Sciences, 12(17): 8467. https://doi.org/10.3390/app12178467

[30] Bakr, M., Abdel-Gaber, S., Nasr, M., Hazman, M.

(2022). Tomato disease detection model based on densenet and transfer learning. Applied Computer Science, 18(2): 56-70. https://doi.org/10.35784/acs-2022-13

[31] Memon, E.D.M.S., Qabulio, M., Kumar, P., Soomro, A.K., Memon, S. (2024). Identification of leaf diseases of different crops using modified ResNet50. The Asian Bulletin of Big Data Management, 4(2): 209-224. https://doi.org/10.62019/abbdm.v4i02.166

[32] Tabbakh, A., Barpanda, S.S. (2023). A deep features extraction model based on the transfer learning model and vision transformer "TLMViT" for plant disease classification. IEEE Access, 11: 45377-45392. https://doi.org/10.1109/ACCESS.2023.3273317

[33] Mandava, M., Vinta, S.R., Ghosh, H., Rahat, I.S. (2024). Identification and categorization of yellow rust infection in wheat through deep learning techniques. EAI Endorsed Transactions on Internet of Things, 10. https://doi.org/10.4108/eetiot.4603

[34] Dermawan, B.A., Awalia, N., Suharso, A., Masruriyah, A.F.N. (2024). The identification of early blight disease on tomato leaves utilizing DenseNet based on transfer learning. E3S Web of Conferences, 500: 01003. https://doi.org/10.1051/e3sconf/202450001003

[35] Islam, M.S Sultana, S., Farid, F.A., Islam, M.N., Rashid, M., Bari, B.S., Hashim, N., Husen, M.N. (2022). Multimodal hybrid deep learning approach to detect tomato leaf disease using attention based dilated convolution feature extractor with logistic regression classification. Sensors, 22(16): 6079. https://doi.org/10.3390/s22166079

[36] Mazumder, M.K.A., Mridha, M.F., Alfarhood, S., Safran, M., Abdullah-Al-Jubair, M., Che, D. (2024). A robust and light-weight transfer learning-based architecture for accurate detection of leaf diseases across multiple plants using less amount of images. Frontiers in Plant Science, 14: 1321877. https://doi.org/10.3389/fpls.2023.1321877

[37] Islam, S. Reza, M.N., Ahmed, S., Samsuzzaman, Cho, Y.J., Noh, D.H., Chung, S.O. (2024). Seedling growth stress quantification based on environmental factors using sensor fusion and image processing. Horticulturae, 10(2): 186. https://doi.org/10.3390/horticulturae10020186

[38] Gehlot, M., Gandhi, G.C. (2023). Design and analysis of tomato leaf disease identification system using improved lightweight customized deep convolutional neural network. In 2023 9th International Conference on Advanced Computing and Communication Systems (ICACCS), Coimbatore, India, pp. 509-516. https://doi.org/10.1109/ICACCS57279.2023.10112920

[39] Divyanth, L.G., Soni, P., Machavaram, R. (2021). Cassava leaf disease dataset. Mendeley Data, V1. https://doi.org/10.17632/3832tx2cb2.1

[40] Getachew, H. (2021). Wheat leaf dataset. Mendeley Data, V1. https://doi.org/10.17632/wgd66f8n6h.1

[41] Mohanty, S.P., Hughes, D.P., Salathé, M. (2016). Using deep learning for image-based plant disease detection. Frontiers in Plant Science, 7: 1419. https://doi.org/10.3389/fpls.2016.01419

[42] Saponara, S., Elhanashi, A. (2022). Impact of image resizing on deep learning detectors for training time and model performance. In Applications in Electronics Pervading Industry, Environment and Society. ApplePies 2021. Lecture Notes in Electrical Engineering, pp. 10-17. https://doi.org/10.1007/978-3-030-95498-7_2

[43] Santoso, C.B., Singadji, M., Purnama, D.G., Abdel, S., Kharismawardani, A. (2024). Enhancing apple leaf disease detection with deep learning: From model training to android app integration. Journal of Applied Data Sciences, 6(1): 377-390. https://doi.org/10.47738/jads.v6i1.507

[44] Sharma, P., Berwal, Y.P.S., Ghai, W. (2019). Enhancement of plant disease detection framework using cloud computing and GPU computing. International Journal of Engineering and Advanced Technology, 9(1): 3139-3141. https://doi.org/10.35940/ijeat.A9541.109119

[45] Karthikeyan, S., Charan, R., Narayanan, S., Anbarasi, L.J. (2025). Enhanced plant disease classification with attention-based convolutional neural network using squeeze and excitation mechanism. Frontiers in Artificial Intelligence, 8: 1640549. https://doi.org/10.3389/frai.2025.1640549

[46] Dhanya, R., Mythili, S., Rexeena, X., Devadas, D. (2025). Detection of tomato leaf diseases using deep learning and spatial attention mechanisms. International Journal of Research and Scientific Innovation, 12(7): 1144-1153. https://doi.org/10.51244/ijrsi.2025.120700118

[47] Althuniyan, N., Al-Shamasneh, A.R., Bawazir, A., Mohiuddin, Z., Bawazir, S. (2024). DeepLeaf: Automated leaf classification using convolutional neural networks. European Scientific Journal ESJ, 20(30): 22. https://doi.org/10.19044/esj.2024.v20n30p22

[48] Venkatesh, T., Ramesh, D. (2025). Integrating spatial and channel features for multi-disease classification using attention-based CNN. Journal of Information Systems Engineering and Management, 10(54s): 181-193. https://doi.org/10.52783/jisem.v10i54s.11050

[49] Amini, S.M., Abbasi-Moghadam, D., Sharifi, A. (2025). Classification of tomato plant leaf disease with entropy filter and convolutional neural network. CABI Agriculture and Bioscience, 6(1): 0050. https://doi.org/10.1079/ab.2025.0050

[50] Kanakala, S., Ningappa, S. (2025). Detection and classification of diseases in multi-crop leaves using LSTM and CNN models. Journal of Innovative Image Processing, 7(1): 161-181. https://doi.org/10.36548/jiip.2025.1.008

[51] Dutta, M. Sujan, M.R.I., Mojumdar, M.U., Chakraborty, N.R., Marouf, A.A., Rokne, J.G., Alhajj, R. (2024). Rice leaf disease classification—A comparative approach using convolutional neural network (CNN), cascading autoencoder with attention residual U-Net (CAAR-U-Net), and MobileNet-V2 architectures. Technologies, 12(11): 214. https://doi.org/10.3390/technologies12110214

[52] Bheemalli, N.S., Kulkarni, B., Aparna, H.D., Swapna, M. (2025). Early and accurate detection of apple leaf diseases using attention U-Net and transformers. International Journal of Environmental Science, 11(21S): 4413-4422. https://doi.org/10.64252/8068db27

[53] Kanimozhi, T., Janakiraman, M., Poomani, M., Jayalakshmi, V. (2024). 3A comparative study of pre-trained transfer learning models in convolutional neural networks for the prediction of diseases in plant leaves. The Bioscan, 19(3): 213-218.

https://doi.org/10.63001/tbs.2024.v19.i03.pp213-218

[54] Mazumder, M.K.A., Kabir, M.M., Rahman, A., Abdullah-Al-Jubair, M., Mridha, M.F. (2024). DenseNet201Plus: Cost-effective transfer-learning architecture for rapid leaf disease identification with attention mechanisms. Heliyon, 10(15): e35625. https://doi.org/10.1016/j.heliyon.2024.e35625