Pak. J. Bot., 58(11): November 2026.

ARTICLE DOI: http://dx.doi.org/10.30848/PJB2026-11(11)

CC-MOBILEVIT: A LIGHTWEIGHT MODEL WITH MOBILE APPLICATION FOR APPLE LEAF DISEASE AUTOMATED RECOGNITION

MINGJIAN ZHANG*, PENGLIANG ZHU, RU TANG, CAN XU AND XIONG LI

Department of Information Technology, Hunan Police Academy, Changsha 410138, China

*Corresponding author's email: mingjianzhang@163.com

Abstract

The cultivation of apple trees plays an irreplaceable strategic role within the world's agricultural system. To mitigate the damaging effects of apple leaf diseases on crop production and food security, it is imperative to implement deep learning (DL)-driven automated systems for early diagnosis. Conventional DL models generally demand designing deep and extensive structures aimed at improving feature extraction performance, but this will lead to the cost of computational overhead. To address this problem, a lightweight leaf disease recognition model called CompConv-MobileViT (CC-MobileViT) is proposed in this paper. CC-MobileViT is an enhanced version of MobileViT, in which the standard convolution is superseded by the CompConv. Compared to MobileViT, the CC-MobileViT model achieves a 50% reduction in parameter count while maintaining the high detection accuracy of 98.10%, which outperforms the state-of-the-art DL models. Moreover, the model has been deployed on the mobile-based WeChat Mini Program, which provides farmers and researchers with the convenience of capturing or uploading images from their smartphones for real-time detection. Therefore, the proposed model strikes a favorable balance between the efficiency and accuracy of disease diagnosis.

Key words: Apple leaf disease detection; CC-MobileViT; Deep learning; CompConv; WeChat Mini Program

Introduction

Apple cultivation constitutes a crucial element of the global agricultural sector, offering farmers across the globe a significant source of income and offering consumers a diverse range of apple products. As a result, apple cultivation is widespread and boasts substantial production volumes, making it a crucial component of the global fruit supply.

Nevertheless, the sustainable development of agriculture faces some major challenges, including the rapid growth of the human population, the increasing demand for agricultural land and resources, and the plant diseases. According to FAO estimates, global crop production suffers annual losses of up to 50% due to plant pests and diseases (Doutoum & Tugrul, 2025). Scab, one of the typical apple leaf diseases, can lead to yield loss of 70% or even more if not managed effectively (Praba *et al.*, 2021).

Most of apple leaf diseases are caused by fungi or viruses, exhibiting characteristics of rapid transmission and widespread detrimental effects. The following is a brief introduction to typical apple leaf diseases, i.e., Alternaria blotch, Brown spot, Grey spot, Mosaic, and Rust.

- *Alternaria* blotch, caused by *Alternaria* fungi, initially forms small brown dots on new apple leaves that later expand into circular lesions with concentric rings, eventually turning gray-white with black specks in the center.
- Apple brown spot is a common, highly harmful apple disease caused by fungi *Diplocarpon mali* and *Marssonina mali*, with three lesion types and prone to massive diseased leaf shedding after wind and rain.
- Apple grey spot, caused by *Phyllosticta pirina* Sacc., develops rapidly in warm, rainy conditions, forming irregular coalesced lesions that lead to leaf withering and scorching.
- Apple mosaic disease, caused by several viruses including apple mosaic virus, leads to leaf discoloration, necrosis, distortion, and premature abscission in severe cases.
- Apple rust, caused by *Gymnosporangium yamadai* Miyabe, primarily infects leaves and young tissues, producing small orange-yellow spots that expand into light-yellow lesions.

In order to increase production capacity and improve the quality of apples, it is important to detect and recognize apple plant diseases accurately. In large-scale orchards, traditional manual inspection of apple leaf diseases entails high labor and time costs. Early symptoms of diseases are often misjudged or overlooked due to variations in the professional competence of inspectors, resulting in delayed disease detection. Furthermore, the limited frequency of manual monitoring makes real-time surveillance challenging, leading to missed optimal control periods and impacting tree health. Hence, it is essential to develop artificial intelligence (AI)-based techniques for automatic apple plant disease detection and recognition.

DL is a method, which employs multi-layer neural network architectures to autonomously learn discriminative features. Convolutional neural network (CNN), on the other hand, is a representative architecture within the realm of DL, which is widely applied in image processing and computer vision tasks (Wang *et al.*, 2020). The prototype of the convolutional neural network can be traced back to the LeNet-5 network in the late 1990s. In 2012, the success of AlexNet (Krizhevsky *et al.*, 2012) in the ImageNet competition propelled CNN to a new level. Subsequently, deep learning network architectures such as VGGNet (Simonyan & Zisserman, 2014), GoogLeNet (Szegedy *et al.*, 2015), DenseNet (Huang *et al.*, 2016), and ResNet (He *et al.*, 2016) emerged one after another, further enhancing the

Received: 01-12-2025 Revised: 12-03-2026 Accepted: 20-03-2026 Online: 10-06-2026

accuracy of image recognition. However, these deep learning neural networks are characterized by their complex architectures, demanding large-scale parameters and computational resources for both training and inference processes. Therefore, researchers have recently explored model light weighting techniques.

Lightweight deep learning neural networks are a type of neural network architecture specifically designed for efficient operation. Compared with traditional large-scale networks, they significantly reduce the number of model parameters and the computational load, while maintaining relatively high performance. This characteristic enables them to operate on resource-constrained devices, such as mobile terminals and embedded systems.

The classical lightweight models include PPLCNet (Cui *et al.*, 2019), ShuffleNetV1-V2 (Zhang *et al.*, 2018; Ma *et al.*, 2018), EfficientNetV1-V2 (Tan & Le, 2019; Tan & Le, 2021), MobileNetV1-V3 (Howard *et al.*, 2017; Sandler *et al.*, 2018; Howard *et al.*, 2019), etc. They not only meet the real-time requirements but also effectively save energy, thus promoting the application of artificial intelligence technology in a wider range of scenarios.

Although lightweight models such as MobileNet and ShuffleNet use hardware-friendly operations to speed up edge inference, they are prone to lower representation capacity. EfficientNet improves efficiency-accuracy trade-offs by means of neural architecture search but still encounters accuracy drops in its compact form for mobile devices.

The principal contributions of this paper are synthesized as follows:

- Conventional DL models for leaf disease recognition face high computational demands due to their deep architectures. This paper presents CC-MobileViT, a lightweight model designed to mitigate this issue. Based on MobileViT, CC-MobileViT incorporates CompConv as a key enhancement over standard convolution with automatic selection mechanism, effectively reducing computational complexity without compromising recognition accuracy. The integration of CompConv enhances the feature extraction ability of MobileViT while maintaining lightweight efficiency, which is rarely explored in existing apple disease recognition models.
- To enhance accessibility, the model is deployed as a WeChat Mini Program. This provides farmers and re-searchers with a convenient tool for real-time leaf disease detection using smartphone cameras, offering a privacy-conscious and offline solution that does not rely on a stable network connection.
- CC-MobileViT outperforms a series of widely used DL models regarding accuracy and parameter efficiency. In comparison to the baseline MobileViT, the proposed model realizes a 50% cut in parameter count while achieving an impressive detection accuracy of 98.10%. This performance trade-off is particularly critical for the tasks of apple leaf disease diagnosis, as the parameter compression significantly alleviates computational and memory burdens without compromising diagnostic reliability. Consequently, the proposed lightweight model enables fast and accurate detection of apple leaf diseases on mobile devices.

Materials and Methods

A. Introduction of the dataset: The dataset consists of 6 categories of disease-related apple leaf images, including Alternaria blotch, Brown spot, Grey spot, Health, Mosaic, and Rust, as shown in Fig. 1. The apple tree leaf images of category Health are sourced from the Plant Pathology 2021-FGVC8 apple tree leaf surface disease dataset released in the 2021 Kaggle Plant Pathology Challenge. The images of other 5 categories are obtained from Paddle Paddle AI Studio development platform (PPAISDP), whose website URL is https://aistudio.baidu.com/. The acquisition of these leaf image data was conducted by data annotation experts, who captured images at various stages of leaf maturity, during different times of the day, and using cameras with diverse focal length settings. All these images can reflect the real-world on-site conditions, which enhance dataset diversity and improve model generalization. Both the 2021-FGVC8 and PPAISDP datasets provide high-quality, professionally annotated labels. Table 1 summarizes the image count across different categories of apple foliage diseases. The composite dataset we constructed includes images of both diseased and healthy apple leaves, which is a notable strength. The integration of images from both FGVC8 and PPAISDP effectively enriches dataset diversity and further enhances the generalization capability of the model.

Table 1. Introduction of the dataset.

Category	Number of images	Source
Alternaria blotch	4809	PPAISDP
Brown spot	5090	PPAISDP
Grey spot	4329	PPAISDP
Healthy	4624	FGVC8
Mosaic	4388	PPAISDP
Rust	5125	PPAISDP

To ensure the diversity and sufficiency of the data samples, the scheme of data augmentation through geometric transformations was employed in the image dataset, including vertical flipping, brightness adjustment, saturation adjustment, hazing, and rotation.

The selected data augmentation schemes improve robustness and fine-grained disease recognition in apple leaves by simulating real-world variations in angle, lighting, and field conditions. By increasing the diversity of training samples, data augmentation mitigates overfitting and fosters the learning of generalized feature representations. Meanwhile, it can strengthen the robustness of the model, enabling it to adapt to various complex real-world scenarios. The results for the operations of image data augmentation are illustrated in Fig. 2.

The dataset was divided into training (70%), validation (15%), and test (15%) sets, which are shown in Table 2.

Table 2. Partition of the dataset.

Dataset	Category					
	Alternaria blotch	Brown spot	Grey spot	Health	Mosaic	Rust
Training	3847	4072	3463	3500	3510	4100
Validation	480	509	432	560	438	512
Test	482	509	434	564	440	513

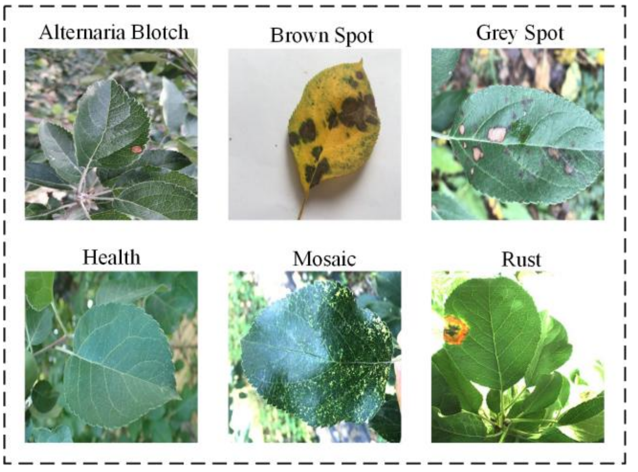

Fig. 1. The leaves of apple trees in six categories.

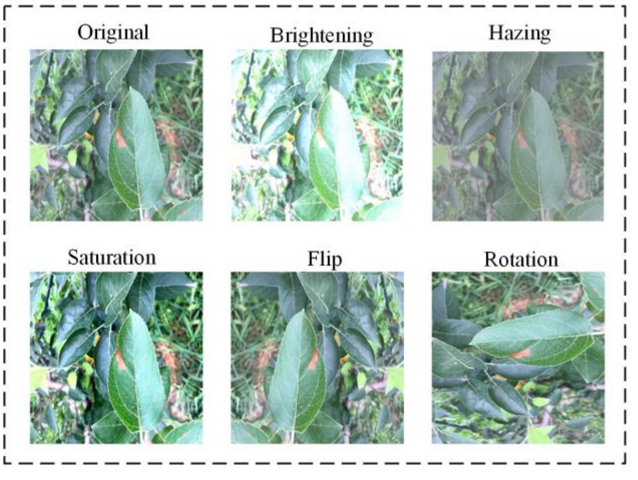

Fig. 2. Operations of image data augmentation.

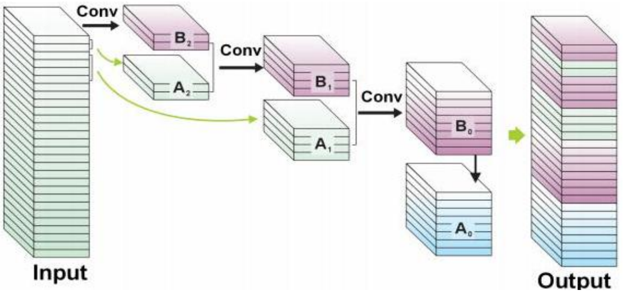

Fig. 3. The structure of CompConv.

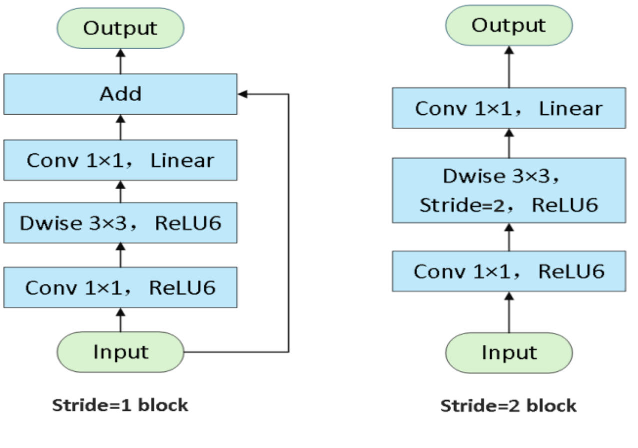

Fig. 4. The inverted residual structure.

B. Proposed CompConv-MobileViT model

(a) Compact convolution: The compact convolution (CompConv) (Zhang *et al*., 2021) is a lightweight convolution module, whose core concept is the spatial processing of partial channels followed by full-channel fusion, which significantly reduces computational load while completely retaining input feature information.

In CC-MobileViT, CompConv is integrated to enhance fine-grained feature preservation while maintaining low architectural complexity.

CompConv employs the divide-and-conquer scheme (see Fig. 3), in which the channel feature map X is formed by the channel-wise concatenation of two independent maps, i.e., X_A and X_B, and then compose them together.

$$X = X_A \oplus WX_B \tag{1}$$

where $\oplus$ represents the channel-wise concatenation operation and W is a learnable parameter for feature map transformation.

Standard convolution (StdConv) requires performing spatial convolution (e.g., 3x3) on all input channels, and the computational overhead and model size are proportional to the number of input channels, resulting in significant redundancy.

In contrast, CompConv uniformly partitions the input channels based on the parameter C_{div}. This strategy designates a fraction of $1/C_{div}$ of the channels to be processed by a spatial convolution layer for local spatial feature extraction. The remaining portions of channels retain their original information without additional spatial convolution computations, thereby reducing the computational overhead of spatial convolution. As a result, CompConv reduces the parameter count by $(1- 1/C_{div})$ times the spatial convolution overhead compared to standard convolution while preserving the completeness of input features. Therefore, it is an efficient convolution solution suitable for resource-constrained scenarios (e.g., mobile devices, edge devices). In this work, C_{div} is configured as 4.

(b) MobileViT: The MobileViT model (Mehta & Rastegari, 2021), as an innovative hybrid architecture, organically integrates CNN and Transformer technologies. Its design fully leverages the spatial inductive bias characteristics of CNN and the significant advantages of Vision Transformer in global feature processing. While achieving model lightweighting, it greatly enhances the efficiency of classification tasks. The core architecture of the MobileViT model consists of MV2 modules, MobileViT blocks, global pooling, standard convolution, and fully connected layer.

The MV2 module, which integrates the inverted residual structure of MobileNetV2, addresses the substantial computational load of traditional convolutions by employing depthwise separable convolutions. The structures of the inverted residual block and MV2 module are illustrated in Figs. 4 and 5, respectively. It is worth noting that, in this paper, the activation function ReLU in the inverted residual structure shown in Fig. 4 is replaced by SiLU. In Fig. 5, DW and PW stand for depthwise convolution and pointwise convolution, respectively.

The MobileViT block, on the other hand, is an innovative optimization of the traditional Vision Transformer (ViT). Its core idea involves splitting the feature map into patches, processing them with Transformer, and then reassembling them. This strategy significantly decreases the computational complexity of the multi-head attention mechanism, effectively improving the model's operational efficiency and achieving lightweighting.

(c) CC-MobileViT: Based on the MobileViT model, the CC-MobileViT model proposed in this paper incorporates a CompConv automatic selection mechanism, as shown in Fig. 6. First, the model adaptively employs convolution kernel dimension standardization to convert convolution kernels of any form into a meta representation (e.g., (3,3)), thereby accurately identifying StdConv operations. This preprocessing step effectively avoids optimization failures caused by ambiguous convolution kernel shapes. Furthermore, when the number of input channels equals that of output channels, the StdConv blocks in the CC-MobileViT model are replaced with the CompConv blocks. The core optimization logic hereby is based on the channel conservation condition, enabling automated selection of CompConv to dynamically optimize computational efficiency. While maintaining feature extraction capability, CC-MoibleViT significantly reduces computational cost and parameter count, which contributes to collaborative optimization of model lightweighting, inference efficiency, and hardware adaptability. This makes it compatible with resource-limited mobile scenarios such as WeChat Mini Program. Compared with existing lightweight ViT and hybrid CNN-Transformer models, CC-MobileViT better balances feature representation and computational efficiency. By incorporating CompConv, it strengthens local feature extraction while preserving the global modeling ability of transformers.

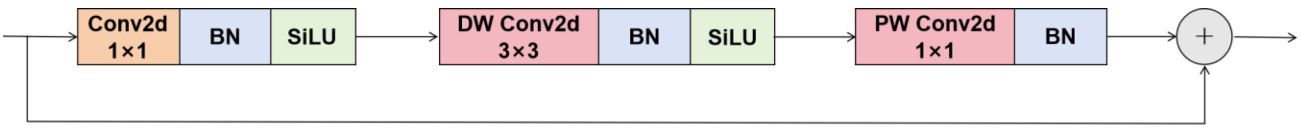

Fig. 5. The structure of MV2 module.

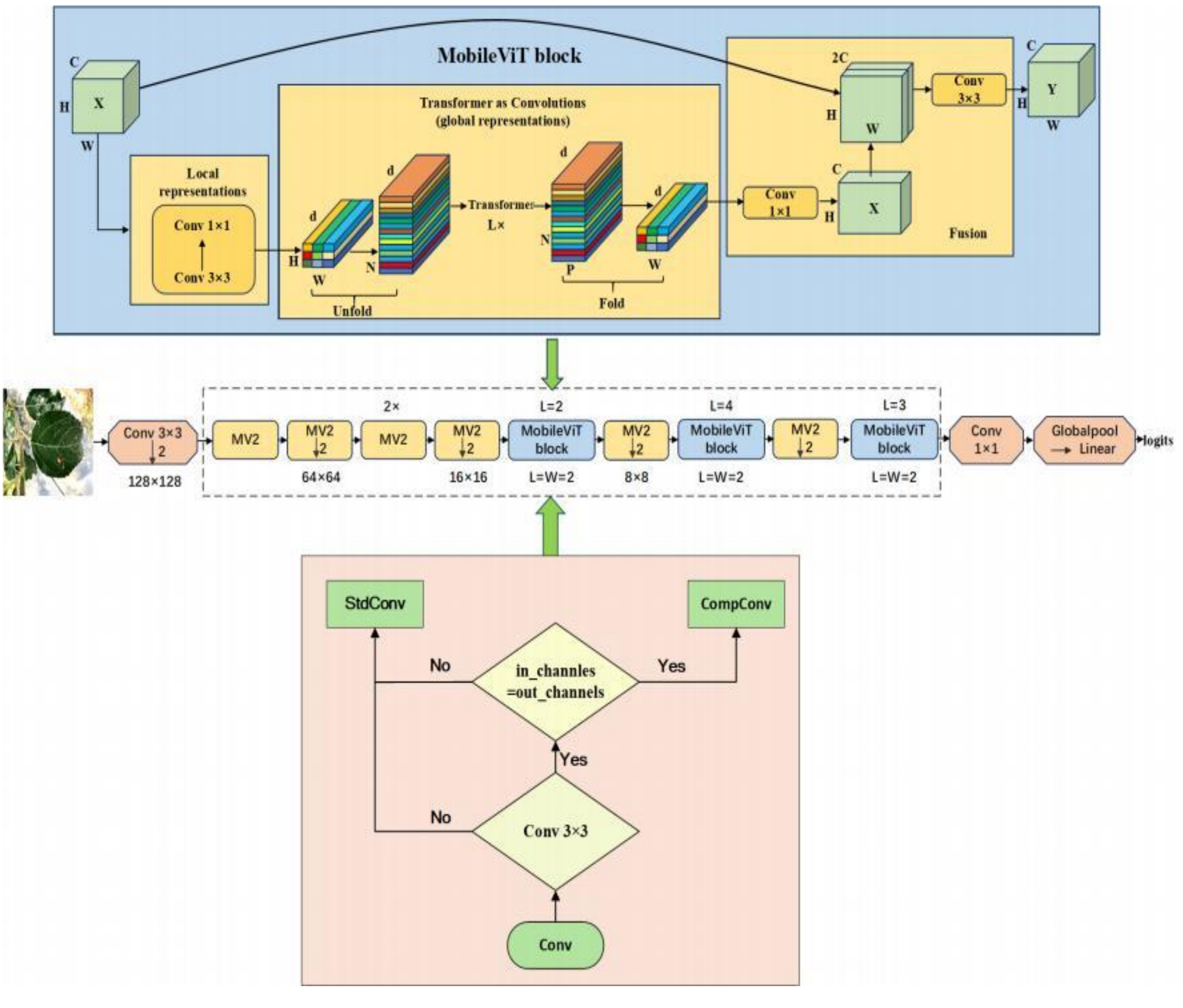

Fig. 6. The improved MobileViT with the CompConv automatic selection mechanism.

Results

A. Establishment of the experimental environment: In this work, we carried out experiments on a remote server. The experimental environmental configuration is presented in Table 3. The model is trained for 50 epochs.

Table 3. Configuration for experiments.

Experimental environment	Item	Configuration
Hardware environment	CPU	AMD EPYC 9754
	GPU	RTX 4090
	Memory	60 GB
	Data Disk	50 GB
Software environment	OS	Ubuntu 22.04
	Version of CUDA	Cuda 12.1
	DL Framework	PyTorch 2.1.2
	Version of Python	Python 3.10

B. Model assessment metrics: When leveraging machine learning models for leaf disease diagnosis, we faced a critical challenge in effectively evaluating a model's diagnostic performance. To this end, four fundamental metrics, i.e., accuracy, precision, recall, and F1-score, which are abbreviated as Acc, Pre, Rec, and F1, respectively, are commonly employed. These scores are rooted in the structure of the confusion matrix, which enumerates correct classifica- tions as true positives (TP) and true negatives (TN), along with incorrect ones as false positives (FP) and false negatives (FN). By conducting an in-depth analysis of the confusion matrix, a more comprehensive understanding of the overall performance and potential weaknesses of the model in handling multi-class disease classification problems can be achieved.

Four model assessment metrics are introduced as follows:

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN} \tag{2}$$

$$\text{Precision} = \frac{TP}{TP + FP} \tag{3}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{4}$$

$$F1 = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{5}$$

Mild class imbalance was observed across the dataset, and we employed Acc, Pre, Rec, and F1 to reliably evaluate classification performance under such imbalance. Four fundamental metrics were adopted as they comprehensively evaluate classification performance and detection reliability in multi-class agricultural disease recognition.

C. Confusion matrix: To evaluate the classification capability of the proposed CC-MobileViT model, a confusion matrix is presented in Fig. 7. As depicted in Fig. 7, our model delivered a high classification accuracy of 98.10% on the test set. Analysis of the confusion matrix revealed that the vast majority of instances in each class were correctly classified. These results confirmed that the model established a clear inter-class separation, effectively mapping instances to their respective categories.

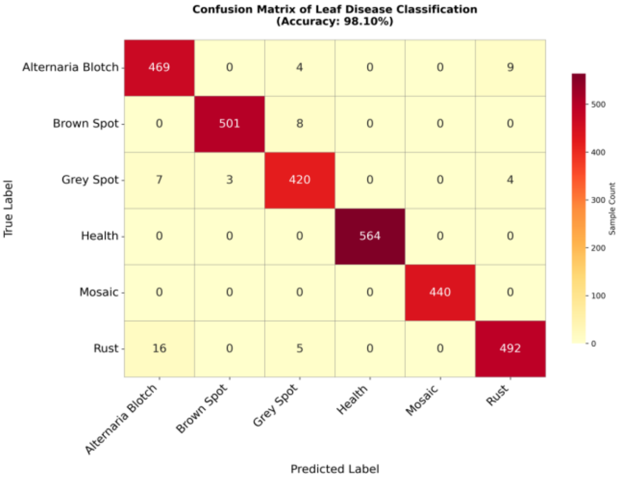

Fig. 7. Detection confusion matrix on CC-MobileViT.

D. Comparison with the classical DL models: To analyze and compare the performance of different classification networks on the leaf disease image dataset, this paper selected several widely used DL models, including AlexNet, DenseNet, VGGNet16, ResNet34, ResNet50, and GoogLeNet, for classification experiments. By carrying out a comparative analysis of the above networks across five metrics, i.e., Acc, Pre, Rec, F1, and parameter count, on the test set, the results are given in Table 4.

From Table 4, it can be seen that the CC-MobileViT model constructed in this study demonstrated better recognition capability on the test set compared to other models. It achieves metrics of 98.10% accuracy, 98.04% precision, 98.07% recall, 98.05% F1, and 0.5 M parameters. Compared to AlexNet, DenseNet, VGGNet16, ResNet34, ResNet50, and GoogLeNet, the accuracy was improved by 2.89, 0.45, 0.79, 1.26, 0.31, and 0.89 percentage points, respectively; precision was enhanced by 2.96, 0.36, 0.79, 1.29, 0.24, and 0.90 percentage points, respectively; recall increased by 3.06, 0.47, 0.82, 1.27, 0.33, and 0.84 percentage points, respectively; F1 rose by 3.03, 0.44, 0.81, 1.30, 0.30, and 0.89 percentage points, respectively; while the parameter count was reduced by 14.1 M, 6.5 M, 133.8 M, 20.8 M, 23.0 M, and 9.8 M, respectively. Due to the strict hardware constraints, the number of parameters (params) is the most critical among the five metrics for mobile deployment in WeChat Mini Program. This validates the distinct advantages of the CC-MobileViT model in handling leaf disease image classification tasks. At the same time, the results show that the potential overfitting risk is effectively alleviated through data augmentation and regularization.

E. Comparison with the classical lightweight models: To analyze and compare the performance of different lightweight DL networks on the leaf disease image dataset, we chose MobileNetV2, EfficientNetV2, ShuffleNet, and MobileViT for comparison with CC-MobileViT. The comparison results are illustrated in Table 5.

Compared to MobileNetV2, EfficientNetV2, and ShuffleNet, the Acc was improved by 5.51, 2.72, and 0.17 percentage points, respectively; Pre was enhanced by 5.61, 2.72, and 0.13 percentage points, respectively; Rec increased by 5.71, 2.89, and 0.18 percentage points, respectively; F1 rose by 5.69, 2.83, and 0.16 percentage points, respectively; while the parameter count was reduced by 1.7 M, 19.7 M and 0.8 M, respectively. MobileViT and CC-MobileViT yielded the same level of accuracy rate. The Pre, Rec, and F1 of CC-MobileViT are very close to those of MobileViT. CC-MobileViT reduces the number of parameters by 50% compared to MobileViT. The performance improvement is attributed to the enhanced ability of the proposed model to capture both local features and global dependencies.

F. Selection of optimization strategies: To investigate the impact of different algorithmic optimization strategies on the performance of the CC-MobileViT model, various optimization measures were systematically combined. The CC-MobileViT model was trained and evaluated accordingly. By calculating and comparing evaluation metrics under each combination, an in-depth analysis of the experimental results was conducted to identify the optimal model configuration. This study systematically explored three dimensions, i.e., optimizer selection, learning rate scheduling strategy, and batch size. The corresponding optimization strategies are provided in Table 6.

In Table 6, small batch size and large batch size are set to 8 and 32, respectively. The fixed learning rate equals to 0.0001. The corresponding performance results are provided in Table 7. The results showed that, when using the AdamW optimization algorithm, a fixed learning rate, and an increased batch size, the CC-MobileViT model achieved the highest accuracy of 98.10%. Additionally, other performance metrics, including Pre, Rec, and F1, achieved optimal results under this configuration.

Table 4. Comparison results of CC-MobileViT and other DL models.

Model	Acc (%)	Pre (%)	Rec (%)	F1 (%)	Params (M)
AlexNet	95.21	95.08	95.01	95.02	14.6
DenseNet	97.65	97.68	97.60	97.61	7.0
VGGNet16	97.31	97.25	97.25	97.24	134.3
ResNet34	96.84	96.75	96.80	96.75	21.3
ResNet50	97.79	97.80	97.74	97.75	23.5
GoogLeNet	97.21	97.14	97.23	97.16	10.3
CC-MobileViT	98.10	98.04	98.07	98.05	0.5

Table 5. Comparison results of CC-MobileViT and other lightweight models.

Model	Acc (%)	Pre (%)	Rec (%)	F1 (%)	Params (M)
MobileNetV2	92.59	92.43	92.36	92.36	2.2
EfficientNetV2	95.38	95.32	95.18	95.22	20.2
ShuffleNet	97.93	97.91	97.89	97.89	1.3
MobileViT	98.10	98.06	98.12	98.07	1.0
CC-MobileViT	98.10	98.04	98.07	98.05	0.5

Table 6. Optimization strategy description.

Optimization strategy	Description
S1	CC-MobileViT + Fixed Learning Rate + RMSProp + Large Batch Size
S2	CC-MobileViT + Fixed Learning Rate + SGD + Large Batch Size
S3	CC-MobileViT + Fixed Learning Rate + AdamW + Small Batch Size
S4	CC-MobileViT + Cosine Annealing Learning Rate + AdamW + Large Batch Size
S5	CC-MobileViT + Fixed Learning Rate + AdamW + Large Batch Size

Table 7. Experiment results for five optimization strategies.

Optimization strategy	Acc (%)	Pre (%)	Rec (%)	F1 (%)
S1	96.50	96.48	96.35	96.39
S2	97.93	97.89	97.91	97.89
S3	97.21	97.16	97.09	97.12
S4	97.72	97.71	97.64	97.66
S5	98.10	98.04	98.07	98.05

G. Wechat mini program development for leaf disease diagnosis: We developed the Apple Leaf Disease Detection System (ALDDS) on the smartphone with the CC-MobileViT model. The integration of the CC-MobileViT model into WeChat Mini Program utilizes a lightweight architecture to mitigate the computation and power consumption constraints of mobile devices. This mobile application facilitates local, real-time AI operation, which not only alleviates concerns regarding data privacy but also removes the need for a persistent network connection. For the deployment of the CC-MobileViT model in a mini-program, the technical pipeline consists of server-side model training, followed by model conversion and optimization to adapt to resource-constrained mobile scenarios.

Upon entering the apple leaf disease recognition interface, users can upload images either by taking photos with their mobile phones or selecting pictures from their albums. The system will intelligently analyze the uploaded images and quickly generate a detailed detection analysis report. The report includes information such as the predicted category, processing time, and the probability of the predicted category matching the true category (TOP1 and TOP2). The intelligent detection results for 6 types of apple leaf disease images on WeChat Mini Program are shown in Fig. 8.

The recognition results of the CC-MobileViT model deployed in WeChat Mini Program offer the probabilities of the top-ranked predicted category matching the true category are 92.70%, 92.65%, 90.35%, 92.77%, 93.59%, and 91.51%, respectively. The inference latency measured on the employed mobile device varies from 267.04 ms to 345.38 ms. This indicates that the CC-MobileViT disease detection system, which is deployed in WeChat Mini Program on mobile phone, is designed to deliver high recognition accuracy and low computational resource consumption simultaneously.

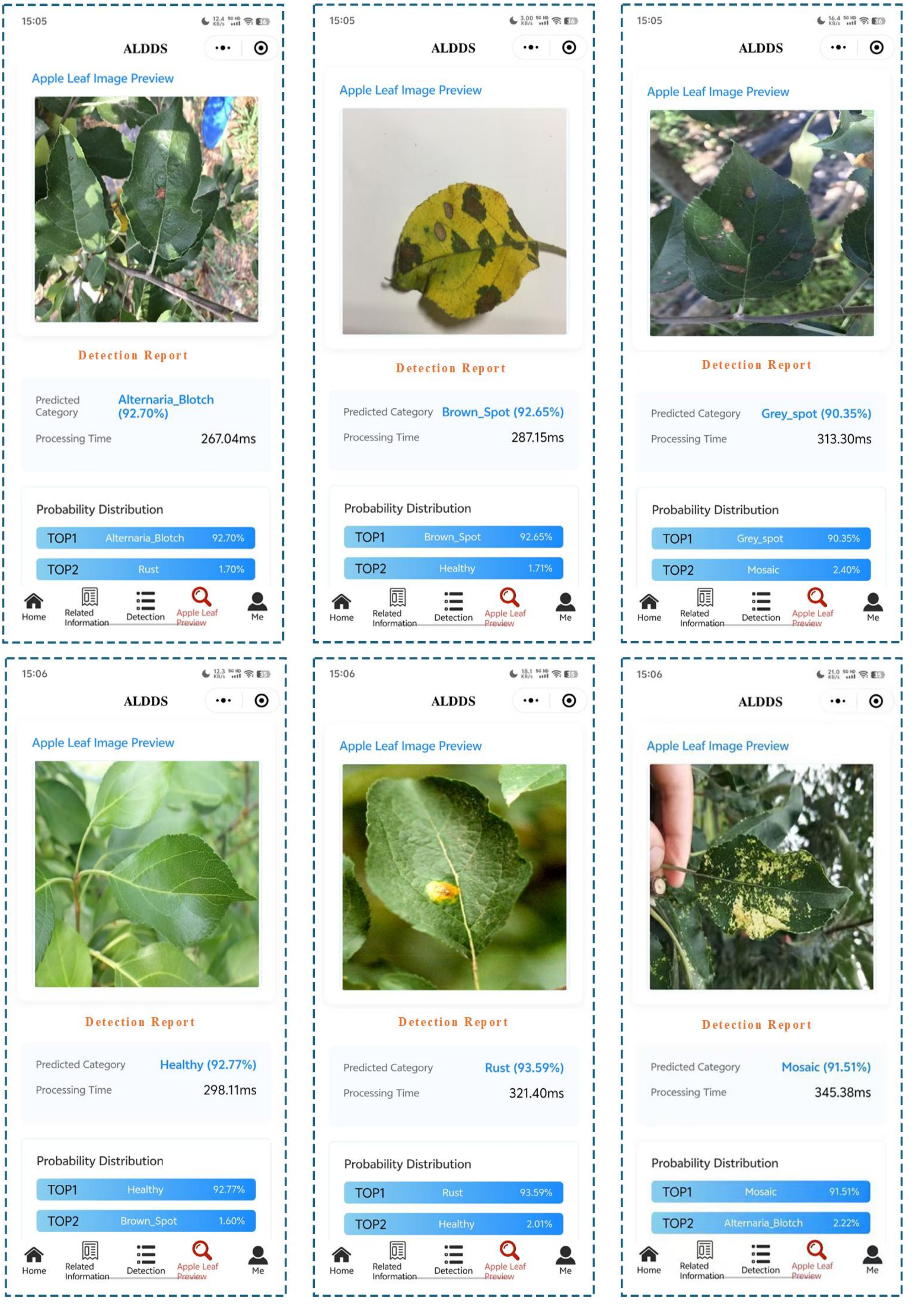

Fig. 8. Intelligent detection results of apple leaf disease images via ALDDS on mobile phone.

Discussion

Over the past decade, an escalating number of scholars have started to employ DL techniques for leaf disease detection purposes (Jiang *et al.*, 2019; Turkoglu *et al.*, 2019; Khan *et al.*, 2022; Tugrul *et al.*, 2022; Gao *et al.*, 2023; Santoso *et al.*, 2024; Pandiyaraju *et al.*, 2025), due to their high accuracy and robust detection capacity.

Deep CNN (DCNN) is a remarkably effective DL method for early disease detection (Sladojevic *et al.*, 2016; Atila *et al.*, 2021; Pandian *et al.*, 2022; Yang *et al.*, 2022; Thanjaivadivel *et al.*, 2025). An improved Faster R-CNN was proposed (Gong & Zhang, 2023) for apple leaf disease detection and obtained an accuracy of 63.1%, where some of the images are from 2021 Plant Pathology Challenge, and the other part is self-created. These DCNN-based approaches usually rely on heavy architectures with high computational costs and limited generalization ability, revealing their inadequacy for real-world agricultural edge deployment.

Notably, hybrid models incorporating multiple heterogeneous frameworks have also been shown to significantly improve detection accuracy in a wide range of practical application scenarios (Irmak & Saygili, 2024). A hybrid deep learning CNN-RNN model was proposed to detect tomato leaf diseases, where the best overall accuracy of 81.75% was obtained for 10-class disease classification (David *et al.*, 2021). CapsNet was presented for the automatic classification and recognition of 10 classes of tomato leaf diseases (Abouelmagd *et al.*, 2024). A deep learning model with complete concatenated block architecture was proposed for multiple crop leaf disease detection on PlantVillage dataset (Arun & Umamaheswari, 2023). A hybrid vision model called APPViT, which combined the features from convolution blocks and multi-head self-attention, was used for plant disease detection (Ullah *et al.*, 2024). Several researchers introduced a hybrid framework for detecting and classifying plant leaf diseases. In this hybrid framework, robust global features were extracted by 3 CNN models including VGG16, Inception-V3, and DenseNet20, while ViT model was employed to extract local features (Aboelenin *et al.*, 2025). A hybrid framework integrating a graph attention network (GAT) branch and a CNN branch achieved remarkable high accuracy in classifying leaf diseases for rice, apple, corn, and wheat (Kerkar *et al.*, 2025). However, these hybrid and structured deep learning methods often suffer from excessive model complexity and insufficient lightweight design. This makes them difficult to deploy efficiently on edge devices for real-time agricultural disease diagnosis.

On the other hand, with the development of image processing and computer vision, the ability to detect and recognize plant leaf diseases using image analysis technique has been remarkably boosted. The application of YOLO enables rapid and precise localization of diseases on leaves, even when they are set against intricate background scenery. By combining object detection model TPH-YOLOV5 and deep learning model MobileNetV3, TPH-YOLOV5-MobileNetV3 network was proposed for multi-objective apple leaf disease (Li *et al.*, 2024). Some research focused on the apple leaf disease detection model by using a combination of improved YOLOv5 network called A-Net and RepVGG module (Liu & Li, 2024). Some investigators proposed the MCDCNet model for detecting five apple leaf diseases using a combination of dual-constrained deformable convolution architecture and a feature integration unit (Liu *et al.*, 2024). MCDCNet achieved an improvement margin of 3.85% relative to the models including YOLO V3, YOLO V7, YOLO V9, etc. Some researchers proposed a target detection model DF-Tiny-YOLO for the identification of apple leaf disease symptoms (Di & Li, 2022).

Although YOLO-based methods achieve promising accuracy in apple leaf disease detection, they suffer from high computational complexity and weak robustness to complex backgrounds, which limits their deployment on edge agricultural devices.

In recent years, significant attention has been paid to the implementation of advanced lightweight models in leaf disease detection by a large number of investigators. A lightweight model called SE-VRNet was designed, which incorporated residual networks and a disease-specific attention mechanism, for leaf disease detection, and achieved top-1 accuracy of 95.71% for self-created leaf disease dataset (Xiao *et al.*, 2023). Some researchers presented a corn leaf disease recognition model based on an improved MobileNetV3-Large architecture (Li *et al.*, 2025). E-AppleNet based on EfficientNetV2, was presented to address the task of apple leaf disease fine-grained discrimination (Banjar *et al.*, 2025). Some researchers proposed a DCNN model with a small number of layers for classification of images from PlantVillage dataset (Vishnoi *et al.*, 2022). An improved EfficientNet network called EfficientNet-MG was proposed to identify apple leaf diseases (Yang *et al.*, 2022). Thanjaivadivel *et al.*, (2025) proposed an enhanced CNN model called EnConv, in which depthwise separable convolution technology and inverted residual module were employed. EnConv obtained an accuracy of 99.87% for 39 categories of plants containing apple, tomato, corn, and potato, etc. Empirical studies have demonstrated that the lightweight models enable rapid, precise identification of apple leaf diseases on resource-limited mobile and edge platforms. Overall, the existing literature demonstrates a consistent pursuit of higher detection accuracy in plant disease diagnosis. However, a critical challenge persists in achieving an optimal balance between model performance and computational efficiency for real-world deployment on edge devices.

Conclusion

A novel lightweight model named CC-MobileViT is developed in this study for apple leaf disease detection. CompConv, a lightweight convolution module, adopts partial-channel spatial processing and full fusion strategy to reduce computation while preserving features for efficient compact extraction. The CC-MobileViT model, which balances both disease diagnostic efficiency and accuracy, enhances MobileViT via substituting its standard convolution with CompConv via automatic selection mechanism. The number of parameters of CC-MobileViT is less than that of MobileViT. The CC-MobileViT leaf disease diagnosis system also offers favorable accuracy rate. In addition, the proposed model has been deployed on the WeChat Mini Program for mobile device application, which enables farmers and researchers to access detection services instantly from their smartphones. This can provide more convenient technical support for smart agricultural production.

The dataset covers 5 common apple leaf diseases, which may lead to insufficient generalization ability when facing complex environments. In the future, we will further improve the generalization ability of the proposed model by validating it on more crop disease datasets and extending it to other plant families for broader agricultural applications.

Author's Contribution: Mingjian Zhang: Conceptualization, methodology, writing—original draft preparation, funding acquisition; Pengliang Zhu: Software, validation, data curation, visualization; Ru Tang: Validation, formal analysis, software, visualization; Can Xu: Resources, software, writing—review and editing; Xiong Li: Formal analysis, data curation, writing—review and editing, validation.

Funding: This work was supported in part by the Key Research and Development Program of Hunan Province, China, under Grant 2025AQ2022; in part by the Scientific Research Fund of Hunan Provincial Department of Education, China, under Grant 25A0695; in part by the Social Science Fund Project of Hunan Province, China, under Grant 22YBA283, and in part by the Ministry of Education Industry-University Cooperation Collaborative Education Project, under Grant 230804978224347.

Conflict of Interests: The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

References

Aboelenin, S., F.A. Elbasheer, M.M. Eltoukhy, W.M. El-Hady and K.M. Hosny. 2025. A hybrid framework for plant leaf disease detection and classification using convolutional neural networks and vision transformer. *Complex Intell. Sys.*, 11: 142.

Abouelmagd, L.M., M.Y. Shams, H.S. Marie and A.E. Hassanien. 2024. An optimized capsule neural networks for tomato leaf disease classification. *EURASIP J. Imag. Video Process*, 2024: 1-21.

Arun, R.A. and S. Umamaheswari. 2023. Effective multi-crop disease detection using pruned complete concatenated deep learning model. *Intell. Syst. Appl.*, 213: 118905.

Atila, U., M. Uçar, K. Akyol and E. Uçar. 2021. Plant leaf disease classification using EfficientNet deep learning model. *Ecol. Inform.*, 61: 101182.

Banjar, A., A. Javed, M. Nawaz and H. Dawood. 2025. E-AppleNet: An enhanced deep learning approach for apple fruit leaf disease classification. *Appl. Fruit Sci.*, 67: 18.

Cui, C., T. Gao, S. Wei, Y.N. Du, R.Y. Guo, S.L. Dong, B. Lu, Y. Zhou, X.Y. Lv and Q.W. Liu. 2019. PP-LCNet: A lightweight CPU convolutional neural network. arXiv: 2109.15099. https://doi.org/10.48550/arXiv.2109.15099

David, H.E., K. Ramalakshmi, R. Venkatesan and G. Hemalatha. 2021. Tomato leaf disease detection using hybrid CNN-RNN model. *Adv. Parallel Comp.*, 38: 593-597.

Di, Jie and Q. Li. 2022. A method of detecting apple leaf diseases based on improved convolutional neural network. *Plos One*, 17: e0262629.

Doutoum, A.S. and B. Tugrul. 2025. A systematic review of deep learning techniques for apple leaf diseases classification and detection. *Peer J. Comp. Sci.*, 11: e2655.

Gao, Y.X., Z.Z. Cao, W.W. Cai, G.F. Gong, G.X. Zhou and L.J. Li. 2023. Apple leaf disease identification in complex background based on BAM-net. *Agronomy*, 13: 1240.

Gong, X.L. and S.J. Zhang. 2023. A high-precision detection method of apple leaf diseases using improved faster R-CNN. *Agriculture*, 13: 240.

He, K.M., X.Y. Zhang, S.Q. Ren and J. Sun. 2016. Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition, 26-30 June, 2016, Las Vegas, Nevada, USA: 770-778.

Howard, A., S. Mark, G. Chu, L.C. Chen, B. Chen, M.X. Tan, W.J. Wang, Y.K. Zhu, R.M. Pang, and V. Vasudevan. 2019. Searching for mobilenetv3. In Proceedings of the IEEE/CVF international conference on computer vision (ICCV), October 27-November 2, 2019, Seoul, Republic of Korea: 1314-1324.

Howard, A.G., M.L. Zhu, B. Chen, D. Kalenichenko, W.J. Wang, T. Weyand, M. Andreetto and H. Adam. 2017. Mobilenets: Efficient convolutional neural networks for mobile vision applications. arXiv:1704.04861. https://doi.org/10.48550/arXiv.1704.04861

Huang, G., Z. Liu, L. Van Der Maaten and K.Q. Weinberger. 2016. Densely connected convolutional networks. arXiv:1608.06993v5. https://doi.org/10.48550/arXiv.1608.06993

Irmak, G. and A. Saygili. 2024. A novel approach for tomato leaf disease classification with deep convolutional neural networks. *J. Agric. Sci.*, 20: 367-385.

Jiang, P., Y.H. Chen, B. Liu, D.J. He and C.Q. Liang. 2019. Real-time detection of apple leaf diseases using deep learning approach based on improved convolutional neural networks. *IEEE Access*, 7: 59069-59080.

Kerkar P.P., T. Veerakumar, A.D. Rahulkar, M.K. Panda and B.N. Subudhi. 2025. A dual-net with graph attention network for plant diseases identification. *IEEE Sens. J.*, 19: 37075-37086.

Khan, A.I., S.M.K. Quadri, S. Banday and J.L. Shah. 2022. Deep diagnosis: A real-time apple leaf disease detection system based on deep learning. *Comp. Electron. Agric.*, 198: 107093.

Krizhevsky, A., I. Sutskever and G.E. Hinton. 2012. Imagenet classification with deep convolutional neural networks. In Advances in neural information processing systems, 3-8 December, Lake Tahoe, Nevada, USA: 1097–1105.

Li, F.M., Y.H. Zheng, S. Liu, F.B. Sun and H.R. Bai. 2024. A multi- objective apple leaf disease detection algorithm based on improved TPH- YOLOV5. *Appl. Fruit Sci.*, 66: 399-415.

Li, H., C. Ruan, J. Zhao, L. Huang, Y. Dong, W. Huang and D. Liang. 2025. Integrating high-frequency detail information for enhanced corn leaf disease recognition: A model utilizing fusion imagery. *Eur. J. Agron.*, 164: 127489.

Liu, B., X.L. Huang, L.M. Sun, X. Wei, Z.Y. Ji and H.X. Zhang. 2024. MCDCNet: Multi-scale constrained deformable convolution network for apple leaf disease detection. *Comput. Electron. Agric.*, 222: 109028.

Liu, Z.Y. and X. Li. 2024. An improved YOLOv5-based apple leaf disease detection method. *Sci. Rep.*, 14: 17508.

Ma, N., X. Zhang, H.T. Zheng and J. Sun. 2018. Shufflenet v2: Practical guidelines for efficient cnn architecture design. In *Proceedings of the European Conference on Computer Vision (ECCV)*, 8-14 September, Munich, Germany: 116-131.

Mehta, S. and M. Rastegari. 2021. Mobilevit: light-weight, general-purpose, and mobile-friendly vision transformer. arXiv: 2110.02178. https://doi.org/10.48550/arXiv.2110.02178

Pandian, J.A., K. Kanchanadevi, V.D. Kumar, E. Jasińska, R. Goňo, Z. Leonowicz and M. Jasiński. 2022. A five convolutional layer deep convolutional neural network for plant leaf disease detection. *Electronics*, 11: 1266.

Pandiyaraju, V., B. Anusha, A.M. Senthil Kumar, K. Jaspin, S. Venkatraman and A. Kannan. 2025. Spatial attention-based hybrid VGG-SVM and VGG-RF frameworks for improved cotton leaf disease detection. *Neural Comp. Appl.*, 37: 8309-8329.

Praba, R.D., R. Vennila, G. Rohini, S. Mithila and K. Kavitha. 2021. Foliar disease classification in apple trees. In International Conference on Advancements in Electrical, Electronics, Communication, Computing and Automation (ICAECA), 8-9 October, Coimbatore, India: 1-5.

Sandler, M., A. Howard, M. Zhu, A. Zhmoginov and L.C. Chen. 2018. Mobilenetv2: Inverted residuals and linear bottlenecks. In Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR), June 18-22, 2018, Salt Lake City, Utah, USA: 4510-4520.

Santoso, C.B., M. Singadji, D.G. Purnama, S. Abdel and A. Kharismawardani. 2024. Enhancing apple leaf disease detection with deep learning: From model training to android app integration. *J. Appl. Data Sci.*, 6: 377-390.

Simonyan, K. and A. Zisserman. 2014. Very deep convolutional networks for large-scale image recognition. arXiv:1409.1556. https://doi.org/10.48550/arXiv.1409.1556

Sladojevic, S., M. Arsenovic, A. Anderla, D. Culibrk and D. Stefanovic. 2016. Deep neural networks based recognition of plant diseases by leaf image classification. *Comput. Intell. Neurosci.*, 2016: 3289801.

Szegedy, C., W. Liu, Y.P. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke and A. Rabinovich. 2015. Going deeper with convolutions. In Proceedings of the IEEE conference on computer vision and pattern recognition, 7-12 June, Boston, Massachusetts, USA: 1-9.

Tan, M. and Q. Le. 2019. Efficientnet: Rethinking model scaling for convolutional neural networks. In International conference on machine learning (ICML), 9-15 June, Long Beach, California, USA: 6105-6114.

Tan, M. and Q. Le. 2021. EfficientNetV2: Smaller models and faster training. In Proceedings of the 38th International Conference on Machine Learning (ICML), 18-24 July, Virtual Event: 10731-10741.

Thanjaivadivel, M., C. Gobinath, J. Vellingiri, S. Kaliraj and J.S. Femilda Josephin. 2025. EnConv: enhanced CNN for leaf disease classification. *J. Plant Dis. Prot.*, 132: 32.

Tugrul, B., E. Elfatimi and R. Eryigit. 2022. Convolutional neural networks in detection of plant leaf diseases: A review. *Agriculture*, 12: 1192.

Turkoglu, M., D. Hanbay and A. Sengur. 2019. Multi-model LSTM-based convolutional neural networks for detection of apple diseases and pests. *J. Amb. Intell. Humaniz Comp.*, 13: 3335-3345.

Ullah, W., K. Javed, M.A. Khan, F.Y. Alghayadh, M.W. Bhatt, I.S. Al Naimi and I. Ofori. 2024. Efficient identification and classification of apple leaf diseases using lightweight vision transformer (ViT). *Discov. Sustain.*, 5: 116.

Vishnoi, V.K., K. Kumar, B. Kumar, S. Mohan and A.A. Khan. 2022. Detection of apple plant diseases using leaf images through convolutional neural network. *IEEE Access*, 11: 6594-6609.

Wang, Q., B. Wu, P. Zhu, P. Li, W. Zuo and Q. Hu. 2020. ECA-Net: Efficient channel attention for deep convolutional neural networks. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 14-19 June, Seattle, Washington, USA (Virtual Conference): 11531-11539.

Xiao, Z.Y., Y.G. Shi, G.L. Zhu, J.P. Xiong and J.H. Wu. 2023. Leaf disease detection based on lightweight deep residual network and attention mechanism. *IEEE Access*, 11: 48248-48258.

Yang, Q., S. Duan and L. Wang. 2022. Efficient identification of apple leaf diseases in the wild using convolutional neural networks. *Agronomy*, 12: 2784.

Zhang, C., Y. Xu and Y. Shen. 2021. CompConv: A compact convolution module for efficient feature learning. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. June 19-25, 2021, Virtual Conference: 3012-3021.

Zhang, X., X. Zhou, M. Lin and J. Sun. 2018. Shufflenet: An extremely efficient convolutional neural network for mobile devices. In Proceedings of the IEEE Conference on Computer Visionand Pattern Recognition (CVPR), 18-22 June, Salt Palace Convention Center, Salt Lake City, Utah, USA: 6848-6856.