



Plant leaf disease identification by parameter-efficient transformer with adapter

Xingshi Xu^{a,b,c}, Guangyuan Yang^{a,b,c}, Yunfei Wang^{a,b,c}, Yuying Shang^{a,b,c},
Zhixin Hua^{a,b,c}, Zheng Wang^{a,b,c}, Huaibo Song^{a,b,c,*}

^a College of Mechanical and Electronic Engineering, Northwest A&F University, Yangling, 712100, Shaanxi, China

^b Shaanxi Key Laboratory of Agricultural Information Perception and Intelligent Service, Yangling, 712100, Shaanxi, China

^c Key Laboratory of Agricultural Internet of Things, Ministry of Agriculture and Rural Affairs, Yangling, 712100, Shaanxi, China

ARTICLE INFO

Keywords:

Plant disease
Adapter
Parameter-efficient
Few-Shot
Precision agriculture

ABSTRACT

Accurate identification of plant disease is of crucial importance for agricultural production. Existing research often tailor dedicated leaf disease recognition models for each plant species to achieve superior recognition accuracy and robustness. However, these approaches are labour-intensive and time-consuming. In addition, when integrating multiple models on a single device, numerous model parameters need to be stored. To address the aforementioned problems, a model named PDNet was designed for the identification of plant leaf disease, advancing the concept of using the same network with almost the same weights for different plant species. First, an Adapter block was utilized to implement a parameter-efficient training. The proposed method could add only 1.39% additional parameters when developing a new model for a specific leaf disease identification task, significantly outperforming the existing 100% fully fine-tuned models. Moreover, to enhance the identification accuracy, Overlapping patch embedding (OPE) was introduced into PDNet to avoid the destruction of essential information on plant leaves. Subsequently, Angular Softmax Loss (A-Softmax Loss) was employed to achieve fine-grained recognition of similar diseases. The proposed method on six plant disease datasets achieved the identification accuracies of 95.16%, 96.74%, 98.26%, 91.16%, 98.67% and 99.58%, respectively. It demonstrated excellent performance under Few-Shot conditions. The recognition accuracy remained above 75% with only three training samples for each disease. This study provides a novel method and paradigm for plant leaf disease identification, offering valuable insights for the field.

Abbreviations used in this research

ViT	vision-transformer
OPE	overlapping patch embedding
A-Softmax	angular softmax
CNN	convolution neural network
FLOPs	the number of floating-point operations
Params	the number of parameters
Tunable-Params	the number of tunable parameters
t-SNE	t-distributed stochastic neighbor embedding
VPT	visual prompt tuning

1. Introduction

Plant diseases affect the overall functioning of plants, reducing the quality and quantity of global agricultural production (Li et al., 2023). Accurate disease identification is critical, as it enables swift preventive measures to curb disease spread, reducing economic losses, enhancing food security and mitigating environmental pollution from fertilizers and pesticides (Thakur et al., 2022; Sachin and Ashish, 2024). However, due to the vast number of plant types and disease categories, along with minimal differences in disease characteristics, manual visual observation is impractical for a precise disease identification. Therefore, accurate and reliable disease identification techniques are of great importance in modern agriculture (Vivek et al., 2022).

Traditional machine vision-based methods for identifying plant leaf

* Corresponding author. College of Mechanical and Electronic Engineering, Northwest A&F University, Yangling, 712100, Shaanxi, China.

E-mail addresses: xingshixu@nwfau.edu.cn (X. Xu), yangguangyuan@nwfau.edu.cn (G. Yang), 15596608508@163.com (Y. Wang), shangyy@nwfau.edu.cn (Y. Shang), hzx@nwfau.edu.cn (Z. Hua), wang.zheng@nwfau.edu.cn (Z. Wang), songhuaibo@nwsuaf.edu.cn (H. Song).

<https://doi.org/10.1016/j.engappai.2024.109466>

Received 30 October 2023; Received in revised form 23 July 2024; Accepted 10 October 2024

Available online 18 October 2024

0952-1976/© 2024 Elsevier Ltd. All rights reserved, including those for text and data mining, AI training, and similar technologies.

diseases rely on manually crafted features. Nevertheless, numerous studies have reported that these methods are insufficiently extractive to distinguish subtle differences in complex disease characteristics (Thakur et al., 2022; Li et al., 2023; Thai et al., 2023). Deep learning has proven effective in disease identification tasks owing to its robust feature extraction capabilities. Initially, some studies treated plant species and disease types as joint categories, using a single deep-learning model for all plant disease identification scenarios (Sladojevic et al., 2016; Sun et al., 2017). However, such approaches increase the number of classes and struggle with distinguishing similar diseases. Moreover, these models need to undergo retraining to satisfy new requirements when plant classes or disease categories increase. Recent efforts have focused on designing specific models tailored to individual plants, such as rice (Sathya and Rajalakshmi, 2022; Chen et al., 2023), tomato (Thangaraj et al., 2021; Kaur et al., 2022), grape (Chen and Wu, 2023; Jin et al., 2023), citrus (Zhang et al., 2022; Vivek et al., 2023a) and peanut (Feng et al., 2023). The primary emphasis of these studies has been on the improvement of models' identification accuracy. These methods identify the disease of each species using different networks with their different weights.

Modelling for each plant species is time-consuming, and the results are difficult to predict due to the variety of plant species and disease types. Recently, Verma et al. (2023) introduced a Meta-learning based framework for recommending models in plant disease identification tasks, enabling the recommendation of the top-n suitable models for an unseen plant. This approach provides a viable alternative to exhaustive model development searches. Nevertheless, in practice, a unified network structure for disease identification across diverse plant species is preferred to streamline model application and deployment (Feng et al., 2023). At present, transfer learning is the dominant and common approach for addressing the aforementioned problems. This method generally requires full fine-tuning of a particular dataset, which identifies the diseases of each plant species using the same network with different weights.

However, despite employing full fine-tuning, maintaining separate model weights for each plant species remains necessary, posing storage challenges when integrating multiple plant disease identification models into devices. Meanwhile, because each model requires a full fine-tuning process, the training of parameters is not efficient and effective enough.

In summary, current methods encounter challenges, such as time-consuming and labour-intensive model design, along with an excessive number of overall model parameters, when aiming to identify leaf diseases across various plant species. To solve the aforementioned problems, in this study, a parameter-efficient vision-transformer (ViT) with Adapter was introduced, named Plant Disease Network (PDNet). Specifically, ViT was utilized as a basic framework owing to its exceptional global context-understanding capability. Then, according to recent research, Adapter is a compact block that could be inserted into the model to adapt its generic features to specific tasks, allowing for efficient fine-tuning and improved performance without extensive retraining. By incorporating an elaborately designed Adapter block, the model achieves parameter-efficient training. The user only needs to add 1.39% additional parameters when developing a new model for a specific leaf disease task, rather than storing an entire set of model parameters. Furthermore, PDNet improved feature extraction with OPE and fine-grained recognition with A-Softmax Loss, thereby enhancing the recognition accuracy of the model. In summary, the proposed method offered a solution that allows the utilization of **the same network with almost the same weights** for each plant species. This approach results in significant savings in storage space and facilitates rapid and efficient training.

The contributions of this research were summarized as listed below.

- (1) A novel plant leaf disease identification model named PDNet was proposed. In which, the Adapter block was utilized to achieve fast and efficient training with fewer parameters. To the best of our

knowledge, this study is the first to develop the same network with almost the same weights for the identification of diseases across different plant species.

- (2) OPE was elaborately designed to avoid the destruction of important information on the plant leaf. A-Softmax Loss was employed to achieve fine-grained recognition of similar diseases by tighter constraints during training.
- (3) Extensive experiments investigated that the proposed model could effectively save storage space and still demonstrate satisfactory identification results under Few-Shot conditions.

2. Related works

The related review was restricted to deep-learning based methods, cognizant that an otherwise important work may be omitted. By reviewing these literature, readers could elucidate the motivation and significance of this research.

2.1. Use a single network to identify leaf disease of many plant species

Previous studies have attempted to use a single network to identify the leaf diseases of various plants, where plant species and their diseases were considered to be joint classes. Sladojevic et al. (2016) applied CaffeNet to disease identification, and the proposed model was able to make a distinction between healthy leaves and 13 different diseases with an overall accuracy of 96.3%. Sun et al. (2017) proposed an improved convolutional neural network (CNN) for disease recognition based on AlexNet, with an average recognition accuracy of 99.56% on the Plantvillage Dataset (14 plant species, a total of 26 different classes of [plant, disease]). Ferentinos (2018) compared the performance of five deep-learning methods for the classification of 58 different classes of [plant, disease] (25 plant species). The results indicated that the visual geometry group (VGG) was the most successful model architecture and that the model identification accuracy reached 99.53%. Nevertheless, these approaches faced a significant challenge when distinguishing subtle variations in leaf disease characteristics among specific plant species and exhibited limitations in their capability to extract the requisite features for achieving precise identification.

2.2. Use different networks to identify leaf disease of different plant species

Considering that the leaf disease characteristics of many plants are extremely similar, researchers designed specialized identification models for each plant species. Sunil et al. (2023) proposed a tomato leaf disease classification model. To enhance the ability of extract disease features, a multi-level feature fusion network and adaptive attention mechanism were employed, and the accuracy reached 99.83%. Ashwini and Sellam (2023) used the Ebola optimization search algorithm to optimize and improve the solution quality of the model while training 3D-DCNN for maize disease, resulting in a commendable model accuracy of 98%. Li et al. (2022) integrated Transformers into CNN, leading to 4.53% improvement in recognition accuracy. This study proved that Transformer could effectively suppress the interference of complex background information on disease recognition. Li et al. (2023) proposed a lightweight multi-branch residual network for tomato leaf disease recognition, which achieved not only high recognition accuracy but also model lightweight (only 4.1M). Kaur et al. (2022) proposed a lightweight Mask R-CNN model for autonomous segmentation and detection of leaf diseases in tomato plants with 98% model accuracy. Vivek et al. (2023b) introduced a collective block and passage layer to avoid the vanishing gradient problem and developed a deep lightweight multi-class classification model for efficient classification of crop leaf disease. Zhou et al. (2023) proposed a residual distilled Transformer model for rice leaf disease identification and achieved satisfactory results with 92% model accuracy. However, these models are typically

tailored to specific plant species rather than for universal application, which often necessitates time-consuming and labour-intensive design processes.

2.3. Use the same network with different weights to identify the leaf disease of different plant species

As it is time-consuming and unpredictable to design specialized leaf disease identification networks for each plant, recent studies have increasingly investigated how to use the same networks for the leaf disease identification of each plant species. The most direct and common method is transfer learning. Long et al. (2018), Xu et al. (2020) and Zhang et al. (2023a) transferred the knowledge of AlexNet, VGG and DenseNet on the ImageNet Dataset to the camellia oleifera, corn and lotus leaf disease identification tasks, respectively, with accuracies of 96.35%, 95.33% and 91.34%. These studies demonstrated the feasibility of transfer learning on disease detection tasks. Meanwhile, another benefit of transfer learning is that it could achieve satisfied results under Few-Shot conditions. Thangaraj et al. (2021) achieved tomato leaf disease detection with an accuracy of 99.55% using transfer learning under Few-Shot conditions. Sudhesh et al. (2023) devoted to analyzing the use of transfer learning to identify rice leaf diseases. They analyzed 10 CNN models and proposed the dynamic mode decomposition approach, which significantly improved the performance of the models. Chen et al. (2023) achieved 95.25% accuracy in rice disease identification using transfer learning with an attention mechanism and domain adaptation networks. Another possible method is multi-task learning. Jiang et al. (2021) employed the alternate learning method to obtain a sharing layer used jointly for recognizing rice and wheat leaf diseases. The accuracy was 97.22% for rice leaf diseases and 98.75% for wheat leaf diseases. However, this method still lacks explorations of disease identification for more plants.

Above all, related works could be summarized as follows.

- (1) **Using a single network** to identify leaf disease of various plants cannot effectively differentiate among similar diseases, and the model needs to be retrained when adding plant species or disease categories;
- (2) **Using different networks** to identify the leaf disease of each plant species requires researchers to dedicate considerable time and effort to design each model;
- (3) **Using the same network with different weights** to identify the leaf disease of different plant species still requires storage space for the entire set of model weights corresponding to each plant.

In this study, the proposed method was designed to identify the leaf disease of different plant species by **using the same network with almost the same weights**. Under the premise of accurately identifying leaf disease, it not only avoids the complicated process of designing a network structure but also effectively saves storage resources for the weights.

3. Datasets and methods

3.1. Dataset description

The dataset used in this research was obtained from several representative datasets, including the PlantVillage Dataset (Hughes and Salathe, 2015), iCassava 2019 Dataset (Ernest et al., 2019), Wheat Leaf Dataset (Hawi, 2021), Rice Leaf Images Datasets (Sethy et al., 2020), Apple Leaf Pathology Dataset (Zhou, 2019) and PSFD-Musa Dataset (Medhi and Deb, 2022). Specifically,

- **The PlantVillage Dataset** includes 14 crop species and consists of 54,303 images of both healthy and unhealthy leaves categorized into 38 groups based on species and disease.

- **The iCassava 2019 Dataset** includes five categories: four prevalent leaf diseases and healthy leaves. All images were captured under in-field conditions by farmers. A total of 5656 labeled images were utilized in the study.
- **The Wheat Leaf Dataset** contains 407 images categorized into three groups: 'Healthy Wheat Leaf', 'Strip Rust' and 'Septoria Disease'. All the images were captured under in-field conditions.
- **The Rice Leaf Images Dataset** is composed of 5932 images captured under in-field conditions. It covers four types of rice leaf diseases: Bacterial Blight, Blast, Brown Spot and Tungro.
- **The Apple Leaf Pathology Dataset** encompasses images of five common apple leaf diseases found in the Loess Plateau region, totaling 26,379 images.
- **The PSFD-Musa Dataset** focuses on banana-related data, comprising 1289 images categorized into three types of leaf diseases.

These datasets covered major cash crops, food crops and vegetable crops as well as fruits, providing strong representation across various plant types. The relevant URLs of the datasets were provided in Appendix A.

To address the presence of identical plant or disease categories across different datasets, a consolidation and reorganization process was developed in this study. Eventually, the integrated dataset contained a total of 93,966 images, the number of plant species was 18, and the number of classes [plant, disease] was 57.

Fig. 1 shows the statistical results of the datasets used in this manuscript, with detailed information provided in Appendix B. As shown in Fig. 1(c), plant diseases have large varieties, and there are numerous plant species whose diseases are extremely difficult to detect. Therefore, constructing a flexible and accurate plant leaf disease identification model is highly necessary. As shown in Fig. 1(b), nearly half of the datasets come from in-field scenarios, where images are often affected by factors such as lighting and shooting distance. These influencing factors can significantly interfere with the model performance.

Sample images from the dataset were shown in Fig. 2. It was worth noting that many of the diseases appeared quite similar to each other (e.g. tomato Early Blight and Late Blight, as shown in Fig. 2(c)). These subtle differences among diseases could be challenging to distinguish through visual observation alone. Consequently, achieving accurate plant leaf disease identification, particularly when aiming to use the same network with almost the same weights, posed significant difficulties.

In this study, 57,597 images from 11 plants (such as apple, banana and corn) were used as a pretraining dataset, which was named D0. This dataset was used to pretrain the initial model and obtain a pretrained model. In addition, datasets for the other six plants (cassava, grape, potato, rice, tomato, and wheat), named D1–D6 respectively, were used for training and testing the final model. The detailed descriptions of each dataset were shown in Table 1.

3.2. Overall architecture

The proposed model architecture in this study was shown in Fig. 3. As illustrated in Fig. 3(a), the proposed model was based on ViT (Dosovitskiy et al., 2020). Contrary to the CNN structure, ViT excels at aggregating long-range contextual information, a capability that has been mainly proven to enhance the recognition of plant diseases (Li et al., 2022; Zhang et al., 2023b).

Different from the vanilla ViT, the proposed method incorporates three key improvements. First, it achieves disease recognition for each plant using a unified network with nearly identical weights. This is facilitated by integrating an Adapter block within the Encoder (as depicted in Fig. 3(b) and (c)) and implementing a parameter-efficient training method. Second, it addresses the feature destruction issue by refining the patch embedding process to better preserve essential information across the image. Finally, the proposed method enhances the

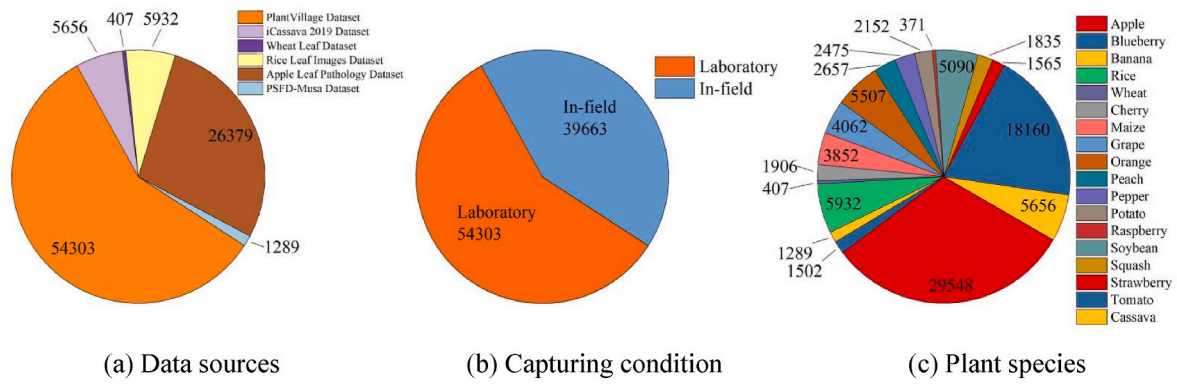


Fig. 1. Statistics of the dataset used in this research.

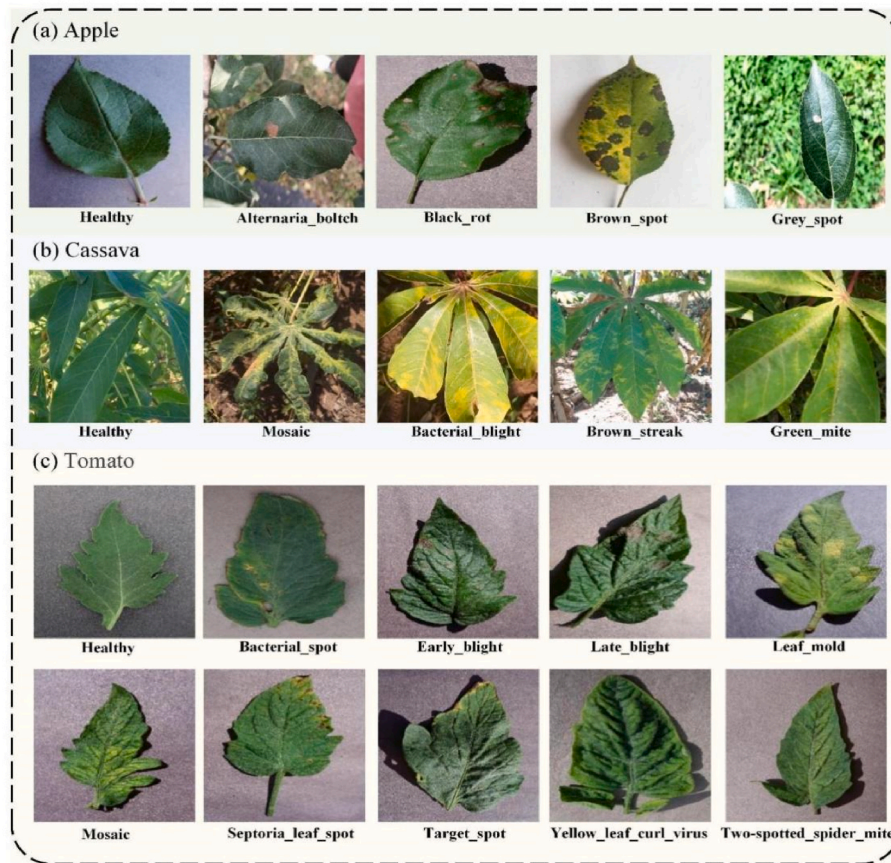


Fig. 2. Examples in the dataset. (a) Apple (captured in the field and laboratory); (b) Cassava (captured in the field); (c) Tomato (captured in laboratory).

Table 1
Description of the dataset used in this study.

Dataset Label	Plants	# Classes	# Images	Application
Dataset0	Apple, Banana, Blueberry, Cherry, Maize, Orange, Peach, Pepper, Raspberry, Soybean, Squash and Strawberry	28	57597	Pretraining
Dataset1	Cassava	5	5656	Training and testing
Dataset2	Grape	4	4062	
Dataset3	Potato	3	2152	
Dataset4	Rice	4	5932	
Dataset5	Tomato	10	18,160	
Dataset6	Wheat	3	407	

ability of the model to distinguish among similar diseases by optimizing the loss function, thereby improving overall classification accuracy and robustness.

3.3. Adapter block

Designing different leaf disease identification models for various plants is time-consuming, laborious and unpredictable (as shown in Fig. 4(a)). Using the full fine-tuning approach to achieve leaf disease recognition for different plants with the same model is inefficient, as it requires saving a full set of weights for each task (as shown in Fig. 4(b)). For the task of leaf disease identification for different plants, the proposed method used the same models with almost the same weights (as shown in Fig. 4(c)).

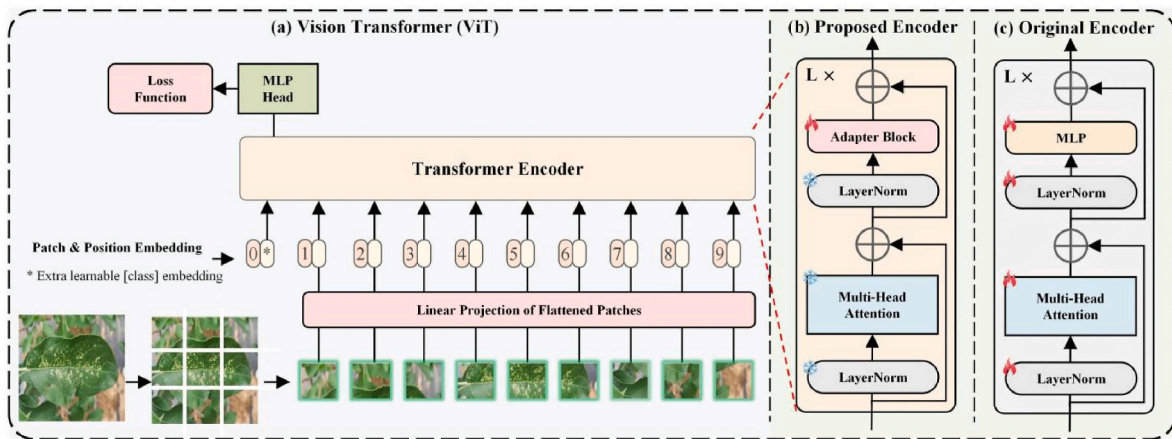


Fig. 3. Architecture of the proposed model. (a) Architecture overview; (b) Proposed Encoder; (c) Original Encoder.

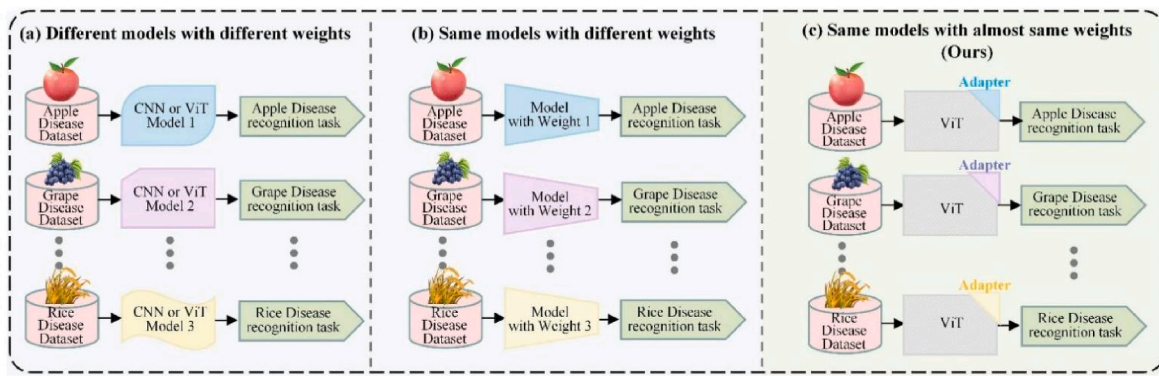


Fig. 4. Differences between the proposed approach and other approaches. (a) Using different models for leaf disease identification in different plants, (b) Using a same models with different weights for leaf disease identification in different plants (full fine-tuning), (c) Using the same models with almost the same weights for leaf disease identification in different plants (ours: Adapter fine-tuning).

The most straightforward approach is to freeze the entire Encoder and only train the last layer of the model. However, this approach results in significantly poorer accuracy than full fine-tuning (Liu et al., 2022). To this end, advanced research has been devoted to parameter-efficient transfer learning, getting a better trade-off between accuracy and low parameters (Gao et al., 2024; Jia et al., 2022; Luo et al., 2023; Xu et al.,

2024b). Adapter is a simple yet powerful module for parameter-efficient transfer learning, which aims to adapt a uniform model to a new task by only changing a few parameters (Houlsby et al., 2019). The limited existing research suggests that this approach holds great potential and strong superiority in terms of accuracy and model size (Chen et al., 2022). As shown in Fig. 4(c), when a new plant needs to be implemented

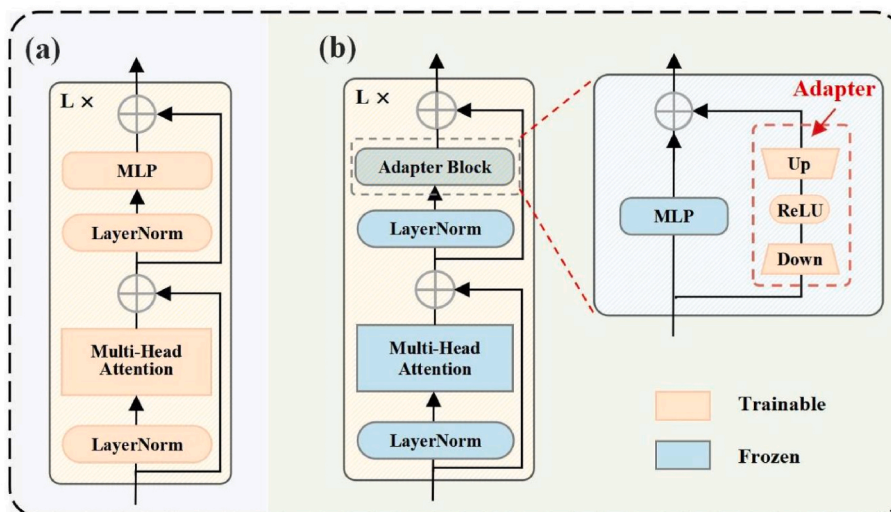


Fig. 5. Comparison of previous full fine-tuning and Adapter fine-tuning. (a) Full fine-tuning; (b) Proposed method.

for leaf disease identification, only the corresponding Adapter needs to be added and saved. This approach greatly reduced the training time of the new model and significantly saved the model storage space.

As shown in Fig. 5(b), the proposed Adapter was inserted in a parallel fashion on the MLP in each Encoder layer. Inspired by Luo et al. (2023) and Chen et al. (2022), the proposed Adapter was designed as a lightweight module with a bottleneck structure. Specifically, it was composed of Down-layer, ReLU and Up-layer. Down-layer projects the features from the d -dimension to the \hat{d} -dimension, and then after non-linear transformation through the ReLU layer, the Up-layer reduced the feature dimension from the \hat{d} -dimension to the d -dimension. In this study, d was set to 32 and \hat{d} was set to 16. Adapter block could be formulated by Eqs. (1) and (2):

$$f(x) = \varphi_{up}(\sigma(\varphi_{down}(x))) \quad (1)$$

$$Adapter(x) = f(x) + MLP(x) \quad (2)$$

where, φ_{up} and φ_{down} denote the down-sampling and up-sampling projections, respectively, σ is the activation function ReLU, and $f(x)$ is the projection feature acquired after Adapter.

The traditional transfer learning method (i.e., full fine-tuning) trains all parameters of the model (as shown in Fig. 5(a)). Contrarily, the proposed method frozen all parts, except the Adapter structure, during fine-tuning (as shown in Fig. 5(b)). In other words, in the whole model, only the parameters in φ_{up} and φ_{down} in Eq. (1) are not frozen. The number of trainable parameters in per layer is only $2 \times d \times \hat{d} + d + \hat{d}$, which includes biases parameters. This approach results in the model being trained with only 1.39% of the parameters of the entire model. With this method, the decoupling of task commonality and task characteristics was realized to a certain extent, the original backbone could focus on the extraction of common characteristics and the Adapter structure could focus on the extraction of characteristic features. This enables the proposed model to effectively and efficiently adapt to various plant disease detection tasks.

3.4. Overlapping patch embedding

The patch embedding operation was improved to avoid vital features of leaf diseases being destroyed. In the ViT work, Dosovitskiy et al. (2020) proposed two approaches for obtaining token embeddings, as shown in Fig. 6(a) and (b), respectively. These two approaches could be interchanged and have a marginal difference effect on the model. The method in Fig. 6(a) was to directly spread each patch first and then map them using fully connected layers. The approach in Fig. 6(b) employed a convolution operation to first achieve the mapping, and then the features were flattened to obtain token embeddings. However, both

approaches shared a common issue where the divided patches did not overlap with each other, resulting in the disruption of local neighbouring structures. This, in turn, led to a decrease in model accuracy, particularly when discriminative regions were inadvertently split (He et al., 2021).

To alleviate this problem, OPE was employed in the proposed model. The improved method, as illustrated in Fig. 6(c), was based on the approach shown in Fig. 6(b). The patch size was increased, allowing neighbouring patches to overlap. In specific realization, the receptive field of convolution was changed from 16×16 to 20×20 while keeping stride equal to 16. In addition, to achieve the same number of patches as the original ViT, padding was set to 2. The proposed approach avoided the destruction of important disease features and enhanced the continuity of local information in the image.

3.5. A-Softmax Loss

A-Softmax Loss was employed to enhance the ability of the model to separate similar diseases. In plant leaf disease identification task, learning robust and discriminative features is essential. The loss function, also known as the objective function, plays a pivotal role in guiding model optimization during training. Some research proved that the most directly and commonly used Softmax Loss may not be always the optimal choice for modelling plant leaf disease recognition (Tavakoli et al., 2021; Pan et al., 2022; Xu et al., 2024a). Original Softmax separates two different classes by a decision boundary. However, in many instances, the disease characteristics of specific plants are extremely similar, and this similarity could lead to a decline in classification performance if intra-class compactness and inter-class separability are not improved.

A-Softmax Loss introduces more stringent decision criteria than Softmax, thereby enhancing the ability of the model to differentiate among similar leaf diseases. To elaborate the theory of A-Softmax Loss, it is essential to revisit the concept of Softmax Loss. The original Softmax Loss converts the initial logits to a probability distribution of [0,1] by the Softmax function and then computes the loss using the cross-entropy loss function. It could be expressed by Eq. (3):

$$L_{\text{Softmax}} = \frac{1}{N} \sum_{i=1}^N -\log p_i = -\frac{1}{N} \sum_{i=1}^N \log \left(\frac{e^{W_j^T x_i + b_j}}{\sum_{j=1}^n e^{W_j^T x_i + b_j}} \right) \quad (3)$$

where, p_i denotes the posterior probability of the i -th image being correctly classified; N is the number of samples involved in the evaluation; and $W_j \in \mathbb{R}^d$ and $b_j \in \mathbb{R}$ are the weights and bias of the j -th class in the last fully connected layer. Considering that bias has a negligible effect, it was omitted for simplicity. Also, as there is an inner product

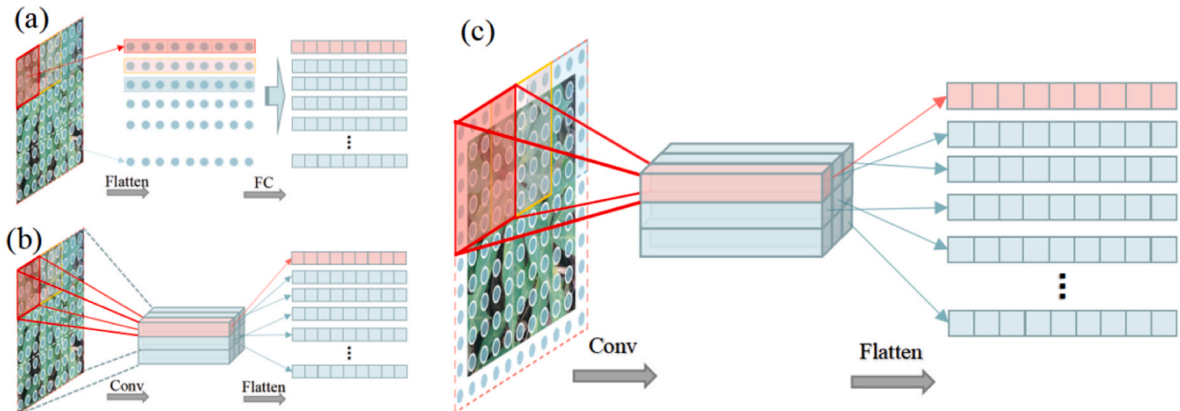


Fig. 6. Schematic of the patch embedding process. (a) and (b) show two routine patch embedding approaches; (c) is overlapping patch embedding.

operation between W and x , $W_j^T x_i + b_j$ can be expressed simply as $\|W_j\| \|x_i\| \cos(\theta_j)$, where θ is the angle between the vectors W and x . Thus, Softmax Loss was expressed as Eq. (4):

$$L_{\text{Softmax}} = \frac{1}{N} \sum_{i=1}^N -\log p_i = -\frac{1}{N} \sum_{i=1}^N \log \left(\frac{e^{\|W_{y_i}\| \|x_i\| \cos(\theta_{y_i})}}{\sum_{j=1}^n e^{\|W_j\| \|x_i\| \cos(\theta_j)}} \right) \quad (4)$$

A-Softmax could be considered as an advanced version of Softmax, which achieves tighter classification by expanding decision boundary. To illustrate, let's consider a binary classification scenario where a sample 'x' belongs to class 1. For the original Softmax, $\|W_1\| \|x\| \cos(\theta_1) > \|W_2\| \|x\| \cos(\theta_2)$ is required to achieve correct classification, whereas A-Softmax changes the decision condition to $\|W_1\| \|x\| \cos(m\theta_1) > \|W_2\| \|x\| \cos(m\theta_2)$ by introducing the hyperparameter m to achieve a tighter classification as Eqs. (5) and (6):

$$L_{\text{A-Softmax}} = \frac{1}{N} \sum_{i=1}^N -\log p_i = -\frac{1}{N} \sum_{i=1}^N \log \left(\frac{e^{\|W_{y_i}\| \|x_i\| \psi(\theta_{y_i})}}{e^{\|W_{y_i}\| \|x_i\| \psi(\theta_{y_i})} + \sum_{j \neq y_i} e^{\|W_j\| \|x_i\| \cos(\theta_j)}} \right) \quad (5)$$

$$\psi(\theta) = \begin{cases} \cos(m\theta), & 0 \leq \theta \leq \frac{\pi}{m} \\ \mathcal{D}(\theta), & \frac{\pi}{m} < \theta \leq \pi \end{cases} \quad (6)$$

where, m denotes a hyperparameter and was set to 3 in this study, d is a monotonically decreasing function, and $\mathcal{D}(\pi/m)$ should be equal to $\cos(\pi/m)$. Referring to previous studies (Liu et al., 2017; Pan et al., 2022), the ψ function used in this study was shown as Eq. (7):

$$\psi(\theta) = (-1)^k \cos(m\theta) - 2k, \theta \in \left[\frac{k\pi}{m}, \frac{(k+1)\pi}{m} \right], k \in [0, m-1] \quad (7)$$

4. Experiments

4.1. Experimental setup

The experiments were conducted using the dataset described in Section 3.1. All datasets were randomly divided into the training set and test set in a ratio of 4:1. At first, DataSet0 was used to obtain a pretrained model, and all parameters of the model were left unfrozen. The best-learned model weights were then used for further training. Subsequently, the pretrained model underwent transfer learning using the appropriate dataset, e.g., Dataset1 was used to train cassava's leaf disease classification model. During the transfer learning process, only the parameters within the Adapter were retrained, and the other weight parameters were frozen.

In the pretraining process (with Dataset0), the SGD optimizer was adopted. The learning rate was set to 0.1, the batch size was 8 and the number of epochs was 100. In the training process (with Dataset1, Dataset2, Dataset3, Dataset4, Dataset5 or Dataset6), considering that the training parameters comprised only 1.39% of the overall model parameters, the learning rate was adjusted downwards to 0.04 to smooth the training. The input size of the images was standardized to $224 \times 224 \times 3$ during all the training and testing.

4.2. Evaluation metrics

To adequately measure the recognition performance, several objective evaluation indices were selected. First, model accuracy was used to measure the evaluation metric of the identification effectiveness of the leaf disease model. It was calculated by Eqs. (8)–(10):

$$\text{Accuracy} = \frac{N_{\text{correct}}}{N_{\text{sum}}} \times 100\% \quad (8)$$

$$N_{\text{correct}} = TP + TN \quad (9)$$

$$N_{\text{sum}} = TP + FP + TN + FN \quad (10)$$

where, TP denotes the number of actual positive samples divided into positive cases, TN represents the number of negative samples correctly predicted, FP represents the number of actual negative cases that are classified as positive, FN represents the number of actual positive cases that are classified as negative.

Furthermore, the number of floating-point operations (FLOPs), the number of parameters (Params), and the number of tunable parameters (Tunable-Params) were selected to comprehensively evaluate the complexity of the model. Specifically, FLOPs was employed to measure the time complexity of the model. Params and Tunable-Params were used to measure the space complexity of the algorithm. It is important to emphasise that as Tunable-Params represents the number of parameters that need to be updated during training, its size directly impacts the speed and efficiency of model training. In general, a smaller Tunable-Params value is desirable because such a network is more flexible, and when adding multiple recognition tasks, the overall model size increases more slowly with a smaller Tunable-Params value.

4.3. Software and hardware platform

In this study, the PyTorch1.7 framework was employed and CUDA11.6 was used as the acceleration environment. The experiments were all conducted under the same conditions, with Ubuntu 18.04.6 LTS as the operating system, Intel(R) Xeon(R) Gold 5217 CPU @ 3.00 GHz as the processor model, Nvidia Tesla V100-PCIE-32GB as the GPU and 128 GB of RAM.

5. Results and discussion

5.1. Experiments on the identification of different diseases

In this study, the proposed model was firstly full fine-tuned using Dataset0 to obtain a pretrained model. Based on this model, Dataset1~6 were used for fine-tuning, which only trained the parameters within the Adapter. Table 2 lists the disease identification results of the proposed model for different plants. As shown in Tables 2 and it is evident that the proposed model consistently achieved an accuracy of over 90% across all datasets.

5.2. Effect of OPE and A-softmax

An ablation experiment was designed to illustrate the practical effectiveness of both OPE and A-Softmax in enhancing the performance of the model. The specific comparison results were shown in Table 3. To ensure a fair comparison, all Adapter fine-tune was based on the pretrained model obtained using both OPE and A-Softmax.

Comparing 1 and 2 (w/o and w/OPE), OPE proved to be valuable for the leaf disease identification task. After using the OPE, the model accuracy improved by 0.21%, 0.58%, 0.67%, 0.53%, 0.70%, 0.05% and 0.23% on the seven datasets, respectively. Fig. 7 shows some gradient-weighted class activation mapping (Grad-CAM) visualisation examples of Dataset5. It could be intuitively seen that the OPE extract features effectively, avoiding destruction of important disease information from plant leaves.

Comparing 1 and 3 (w/o and w/A-Softmax), A-Softmax Loss effectively improved the accuracy of the model. When using A-Softmax Loss instead of Softmax Loss, the identification accuracy improved by 1.31%, 3.00%, 0.48%, 0.10%, 0.01%, 0.03% and 1.19% on each of the seven

Table 2
Detection performance of PDNet on different plants.

Stage	Accuracy (%)						
	Dataset0 (pretrain)	Dataset1 (Cassava)	Dataset2 (Grape)	Dataset3 (Potato)	Dataset4 (Rice)	Dataset5 (Tomato)	Dataset6 (Wheat)
Train	98.64	95.72	97.52	98.44	94.42	99.28	99.95
Test	98.62	95.16	96.74	98.26	91.16	98.67	99.58

Table 3
Ablation experimental results of OPE and A-Softmax.

Num	OPE	A-Softmax	Accuracy (%)						
			Dataset0 (pretrain)	Dataset1 (Cassava)	Dataset2 (Grape)	Dataset3 (Potato)	Dataset4 (Rice)	Dataset5 (Tomato)	Dataset6 (Wheat)
1			97.28	91.94	95.48	97.58	90.21	98.26	98.14
2	✓		97.49	92.52	96.15	98.11	90.91	98.31	98.37
3		✓	98.59	94.94	95.96	97.68	90.22	98.29	99.33
4	✓	✓	98.62	95.16	96.74	98.26	91.16	98.67	99.58

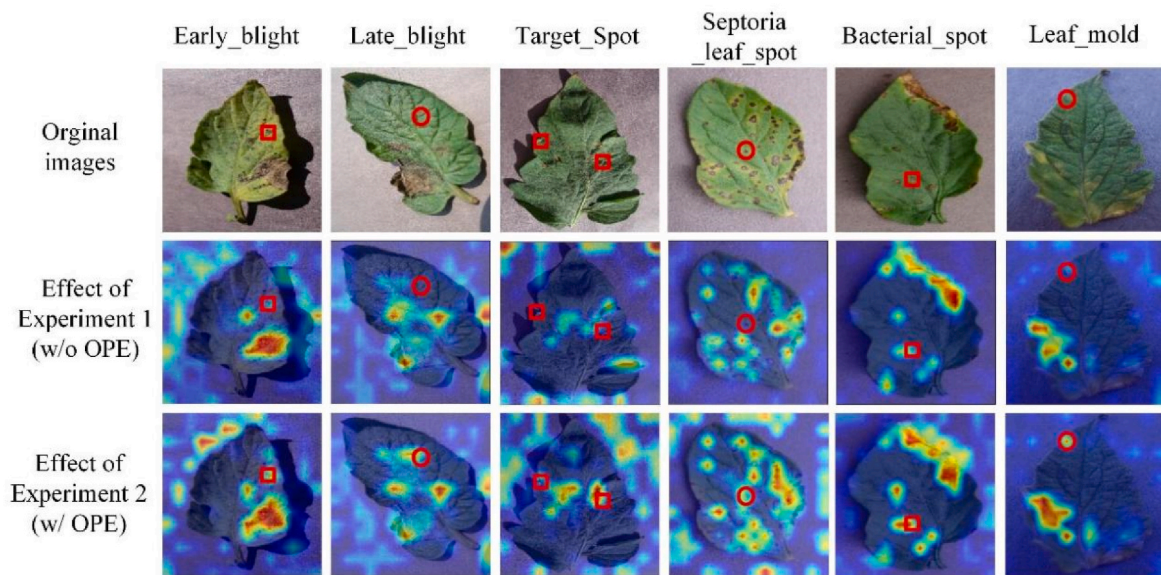


Fig. 7. Grad-CAM visualisation examples.

datasets. In the best-case scenario, this approach resulted in a remarkable 3% boost in accuracy. The t-distributed stochastic neighbor embedding (t-SNE) is an advanced reduced dimensional and visualisation algorithm (Linderman and Steinerberger, 2017). It could map the data from high-dimensional feature space to low-dimensional space to observe the data distribution and the feature extraction effect of the model. In this study, to visually analyse how A-Softmax improved the model performance, Dataset0 was used as an example to visualise the high-dimensional feature distributions obtained after image embedding

using the t-SNE algorithm. As shown in Fig. 8, each point represents an image, and different colours indicate different categories. Fig. 8(a) shows the distribution of embedded features obtained from the un-trained model; Fig. 8(b) shows the distribution of embedded features obtained from the model that has been trained with Softmax loss; and Fig. 8(c) shows the distribution of embedded features obtained from the model that has been trained with A-Softmax loss.

As shown in Fig. 8(a), before training, the distribution of features was dispersed, and the features of different classes were interspersed with

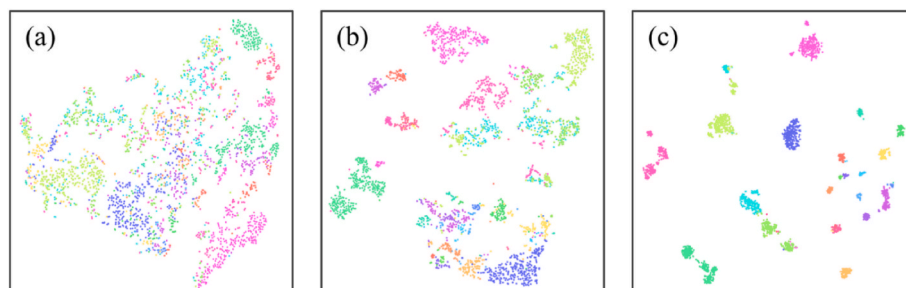


Fig. 8. t-SNE visualisations on Dataset0. (a) Before training; (b) After training with Softmax; (c) After training with A-Softmax (Ours).

each other. As shown in Fig. 8(b), after training with Softmax, the embedded features of different classes were realized to be divided in the space. However, their boundaries between different classes were mostly bordered, and the decision boundary between classes was not sufficiently distinct. When the image was disturbed by lighting, blurring or other factors, the model was prone to identification errors. The embedded features after using A-Softmax were shown in Fig. 8(c), where each class was aggregated together, and there were large boundaries between classes. Consequently, the model could still achieve satisfactory identification even in cases of deteriorated image quality.

5.3. Effect of adapter

To demonstrate the great advantages and utility of Adapter, three different fine-tuning methods were compared, and the comparison results were shown in Table 4. These methods include Full-tuning, which involves fine-tuning all the model parameters based on the pretrained model for a specific task; Linear, which entails freezing the entire feature extraction network and only training the weight parameters of the final fully connected layer; and visual prompt tuning (VPT), which is an emerging and efficient parameter fine-tuning method (Jia et al., 2022).

As shown in Table 4, when using the Full-tuning method for fine-tuning, 85.80 M parameters need to be tuned. Storing the weights of these parameters will take up a lot of storage cost. In particular, when one device needs to store the weights of multiple disease recognition models, the storage cost is not negligible. Linear, VPT (Jia et al., 2022) and the proposed method only need to save an additional weights of 0.07 M, 0.59 M and 1.19 M parameters when training a new disease identification model, which were only 0.08%, 0.69% and 1.39% of the whole model, respectively. The results indicated that although the Linear method required the least number of training parameters, the model accuracy obtained was poor, with a drop in accuracy of more than 10% on some challenging datasets. This was because the frozen backbone was not well adapted to the disease characteristics of specific plants, and some important disease characteristics were not sufficiently extracted. The model accuracy using the VPT method was higher than that using the Linear method, but there was still a gap from using the Full-tuning method. Using the proposed method, the model's recognition accuracy approached or even exceeded that of the full-tuning method while training only 1.19 M parameters. The recognition accuracies on the six datasets were 95.16%, 96.74%, 98.26%, 91.16%, 98.67%, and 99.58%, respectively.

Fig. 9 intuitively illustrates the performance of different methods. The closer each curve intersects with the axes towards the outer edges, the better the metric. The larger the area enclosed by the curve, the better the overall performance of the algorithm. It can be observed that the proposed method demonstrates satisfactory performance in terms of the Tunable-Params and Accuracy across various datasets, meeting the requirements for constructing a flexible and accurate plant leaf disease identification model.

5.4. Results under the Few-Shot condition

Data collection is a time-consuming and labour-intensive process. One advantage of transfer learning is its ability to effectively train models with a small number of labeled samples. Considering that the

Table 4

Different fine-tuning methods with the pretrained model.

Method	Tunable-Params (M)	Accuracy (%)					
		Dataset1 (Cassava)	Dataset2 (Grape)	Dataset3 (Potato)	Dataset4 (Rice)	Dataset5 (Tomato)	Dataset6 (Wheat)
Full-tuning	85.80 (100%)	98.00	96.94	98.19	91.18	98.64	99.61
Linear	0.07 (0.08%)	84.88 (-13.12)	92.89 (-4.05)	86.16 (-12.03)	85.56 (-5.62)	87.51 (-11.13)	97.14 (-2.47)
VPT	0.59 (0.69%)	91.46 (-6.54)	96.70 (-0.24)	97.96 (-0.23)	90.16 (-1.02)	98.11 (-0.53)	99.58 (-0.03)
Ours	1.19 (1.39%)	95.16 (-2.84)	96.74 (-0.20)	98.26 (+0.07)	91.16 (-0.02)	98.67 (+0.03)	99.58 (-0.03)

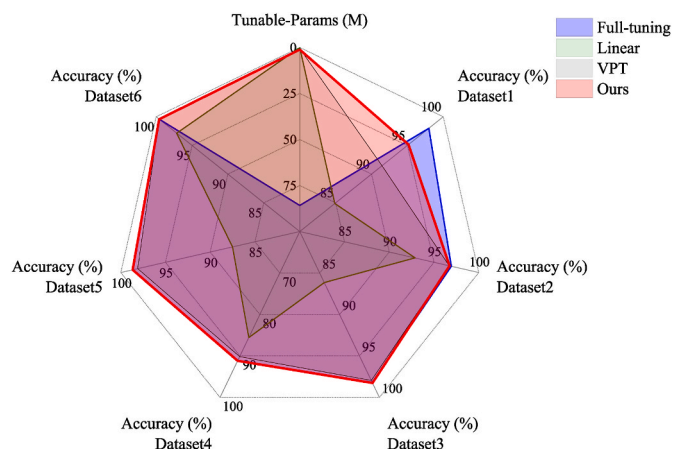


Fig. 9. Comparison results of Tunable-Params and Accuracy of the four methods.

proposed method belongs to the variant of transfer learning, the performance of the proposed method under Few-Shot conditions was explored, and comparison experiments between the proposed method and other methods were conducted. Three other transfer learning methods were compared, where only 10, 5 and 3 images were provided as training samples for each class in each set of experiments, respectively. The models' accuracy under different conditions was shown in Fig. 10.

The accuracy of all methods decreased as the number of training samples decreased. This phenomenon indicates that when provided with a limited amount of training data, it becomes challenging for the model to acquire discriminative and robust features, ultimately impacting its performance. The Full-tuning method typically exhibit low recognition accuracy due to model overfitting when the number of samples is particularly small (e.g., 3 images per class). The Linear method is consistently less accurate, which is attributed to the low feature extraction capability of the model. The VPT method performed better than the Full-tuning and the Linear methods, and was able to avoid serious overfitting and underfitting. Nevertheless, compared with the other three methods, the proposed method achieved the highest recognition accuracy when only 10, 5 and 3 images were provided as training samples for each class, respectively. Even in the case of only 3 training images per class, the proposed method achieved an accuracy exceeding 70%. These results indicated that the proposed method performs satisfactorily under Few-Shot conditions.

5.5. Discussion with existing leaf disease identification works

Using CNNs for plant leaf disease recognition has been a mainstream in recent years (Thakur et al., 2022). The most commonly used baseline models include ResNet50, VGG16 and MobileNet. However, the CNN-based method employed a local receptive field when processing images, which means that each convolutional kernel focused on only a small region and may not effectively capture global information. Transformer models have gained attention for their ability to handle long-term dependencies. Li et al. (2022) demonstrated the effectiveness

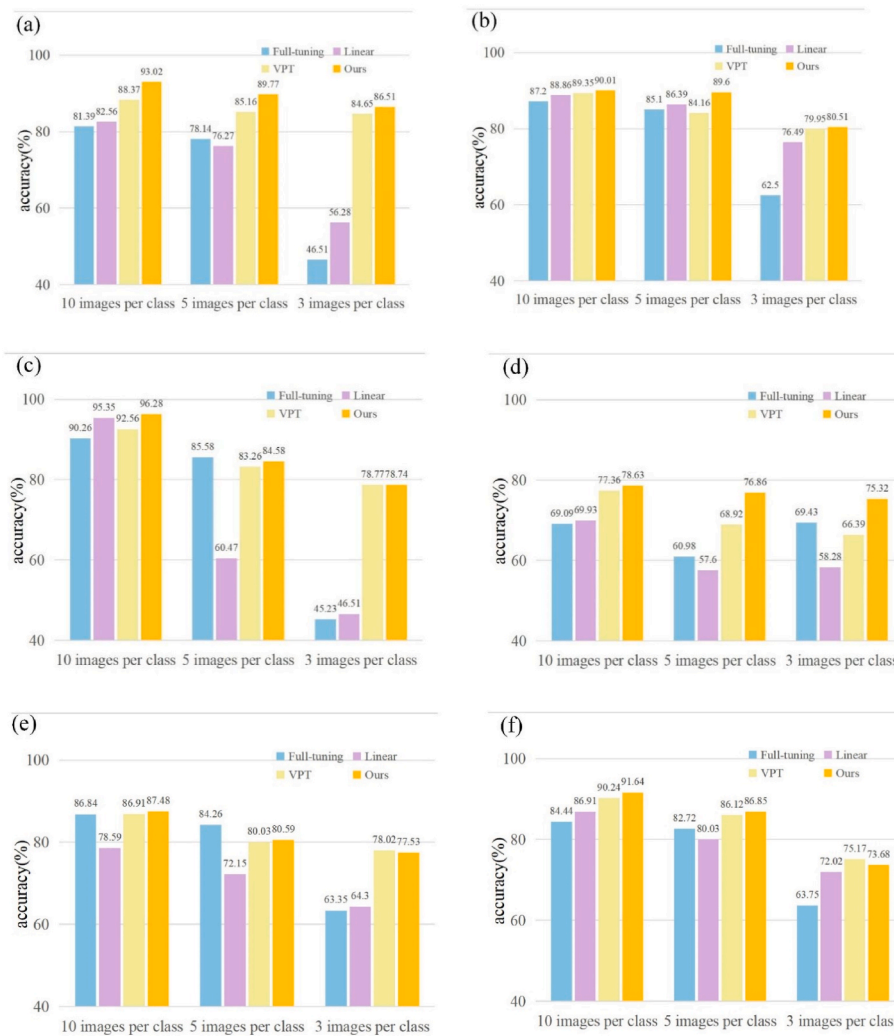


Fig. 10. Results of different methods under Few-Shot conditions. (a), (b), (c), (d), (e) and (f) denote the recognition results on Dataset1, Dataset3, Dataset4, Dataset5, and Dataset6, respectively.

of Transformer in leaf disease recognition task by inserting Transformer into a CNN framework. Recently, many researchers have actively explored how to use ViT elegantly for plant leaf disease recognition (Guo et al., 2022; Thai et al., 2023; Yang et al., 2023; Zhou et al., 2023).

However, the Transformer model has a huge number of parameters compared with most CNN models and is subject to storage constraints when deploying multiple leaf disease recognition models simultaneously. In this study, the proposed model realized the same model with the same weight for the disease recognition task in different plants and provided a feasible solution to save the model storage resource. The proposed model should store 85.80 M parameters initially, and only 1.19 M parameter weights need to be stored additionally when a new model is developed for a new leaf disease task. When using common CNN-based methods to develop a new model, it is necessary to store the parameter weights of an entire model. For example, using ResNet50 as the backbone, developing 5 models results in a total parameter count of up to 128.15 M (the proposed model only needs 90.56 M); developing 10 models results in a total parameter count of up to 256.30 M (the proposed model only needs 96.51 M). It could be seen that the advantages of the proposed method in terms of storage become more obvious when more models are developed.

Model lightweighting is also a current research focus (Li et al., 2023; Xu et al., 2023). Some scholars have developed lightweight leaf disease recognition models with even less than 5 M parameters (Lu et al., 2022;

Li et al., 2023). However, a systematic review by Thakur et al. (2022) showed that these models were developed for individual species and required technicians to design the model structure in a targeted manner. This study also showed that lightweight methods typically fail to achieve high classification accuracy on complex datasets. Contrarily, the proposed method used a unified framework that avoids complex design processes for different species, which was more convenient for both the development and application stages.

5.6. Discussion with existing leaf disease identification works

To further explore the advantages of the adapter method proposed in this study, several common CNN-based models were compared, including ResNet34, ResNet50, VGG16, MobileNetv2 and ShuffleNetv2. It is noteworthy that the adapter technique was not applied to these CNN-based models. This decision stems from the suitability of the Adapter block for integration within the Multi-Head Self-Attention mechanism of Transformer-based models, facilitating effective feature extraction. Contrarily, CNN models have relatively fixed architectures, limiting the flexibility of Adapter block integration and adjustment as observed in ViT. Therefore, for CNN-based models, Full-tuning or Linear methods were employed for fine-tuning in the study, and the comparative results were shown in Table 5.

As shown in Table 5, the proposed model achieves the best

Table 5
Performance of different models using different fine-tuning strategies.

Method	FLOPs (G)	Params (M)	Tunable-Params (M)	Accuracy (%)					
				Dataset1 (Cassava)	Dataset2 (Grape)	Dataset3 (Potato)	Dataset4 (Rice)	Dataset5 (Tomato)	Dataset6 (Wheat)
ResNet34 Full-tuning	3.68	21.79	21.79	83.02	90.35	92.56	89.88	93.06	95.06
ResNet34 Linear			0.07	70.73	82.15	86.05	83.98	89.26	90.12
ResNet50 Full-tuning	4.09	25.63	25.63	84.00	92.71	93.02	90.55	93.36	96.29
ResNet50 Linear			0.07	72.06	83.99	87.44	83.98	89.48	91.36
VGG16 Full-tuning	15.47	138.36	138.36	84.88	93.87	93.72	92.92	93.80	95.06
VGG16 Linear			0.07	71.26	86.21	88.14	89.88	89.65	92.59
MobileNetv2 Full-tuning	0.33	3.50	3.50	83.82	92.10	93.02	92.75	93.75	93.83
MobileNetv2 Linear			0.07	58.62	84.24	84.19	84.65	88.68	86.42
ShuffleNetv2 Full-tuning	0.15	2.28	2.28	84.53	91.23	93.25	90.64	94.08	93.83
ShuffleNetv2 Linear			0.07	60.04	82.27	85.12	87.77	88.44	88.89
Ours	45.35	85.80	1.19	95.16	96.74	98.26	91.16	98.67	99.58

recognition results across all datasets. While some lightweight CNN-based models have lower FLOPs and Params, their recognition performance is often unsatisfactory without a careful design. Designing effective and lightweight models tailored to each plant species requires considerable effort. The proposed method utilizes a single model for all types of leaf disease recognition tasks, simplifying the model design process while maintaining high recognition accuracy. In addition, similar to the Linear method, the proposed approach freezes most parameters during training, with only 1.19 M Tunable-Params, yet achieves satisfactory results. It exhibits flexibility in adapting to new recognition tasks without significant increases in the overall model size.

5.7. Results on generalization experiment

To evaluate the generalization ability of the proposed method, 60 labeled in-field grape leaf images were randomly selected from the Internet to form Dataset7. The four models trained on Dataset2, as described in Section 5.3, were directly tested on Dataset7. Sample

images from Dataset2 and Dataset7 are illustrated in Fig. 11(a) and (b), respectively. It is evident that real-world field images contain additional interfering factors, such as lighting conditions, shooting angles and background environments, which potentially influence the model performance.

Table 6 shows the identification results of different models on Dataset2 and Dataset7. The results indicated that all models exhibited a

Table 6
Results of the generalization experiments.

Method	Accuracy (%)	
	Dataset2	Dataset7
Full-tuning	96.94	88.33
Linear	84.88	70.00
VPT	91.46	86.67
Ours	95.16	91.67

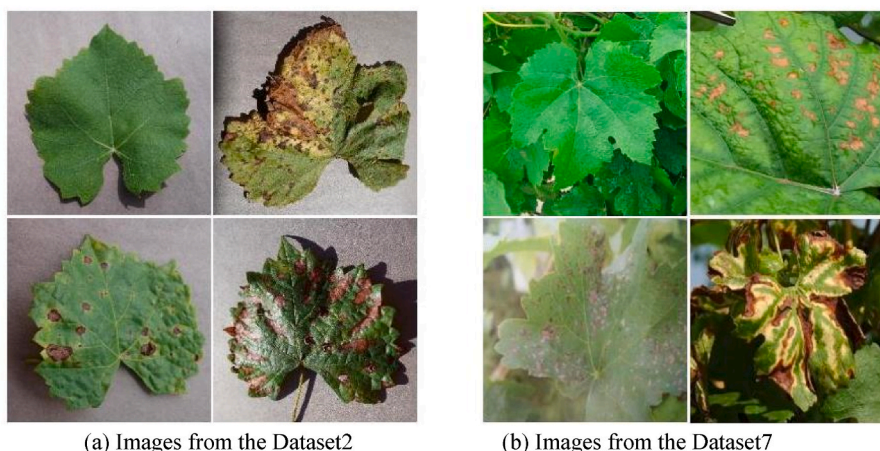


Fig. 11. Grape leaf images from the Dataset2 and the Dataset7.

decline in accuracy on Dataset7. This phenomenon could be attributed to the differences in the data distribution between Dataset2 and Dataset7, as well as the presence of more complex interfering factors in the latter's images. Among the four models, those trained using the Full-tuning and Linear methods showed significant decreases in accuracy. This could stem from overfitting on Dataset2 in the case of Full-tuning, resulting in poor generalization. Conversely, the Linear method may have inadequately extracted discriminative features for grape leaf diseases, leading to lower performance under more challenging conditions. The VPT and the proposed method demonstrated relatively smaller decreases in accuracy. Notably, the proposed method achieved the highest accuracy of 91.67% on Dataset7, highlighting its robust generalization capabilities. Despite varying data distributions and complex interference factors, the proposed method maintains high accuracy in identifying plant leaf diseases.

5.8. Managerial implications

The managerial implications of the proposed model was further discussed in this section. This work presented a novel approach to plant leaf disease identification, considering the model design, deployment, and application aspects of the real world condition.

From the model design perspective, the vast and ever-expanding variety of plants and leaf diseases makes designing different models to cover the entire problem space labor-intensive and inefficient. This study provided a unified framework for the identification of various plant leaf diseases. This framework, through parameter-efficient fine-tuning, effectively replaced the complex model design process, saving computational resources and time during model development. Since the majority of the parameters are frozen, the model performs excellently even with limited sample sizes, which not only enhances the model's generalization ability but also significantly reduces the cost of dataset creation.

From the model deployment perspective, the proposed model employed adapter technology to significantly reduce the total number of parameters, ensuring practicality in resource-constrained environments. This design robustly supports the further deployment and integration of the model into existing agricultural management systems or agricultural equipment such as UAVs and AGVs.

From the model application perspective, farm managers can easily use the model to accurately identify leaf diseases without requiring specialized knowledge. This user-friendly approach enables non-experts to make informed decisions, thereby improving overall farm management efficiency and enhancing the timeliness and accuracy of disease control.

In summary, the model's ease of design, deployment, and application effectively enhances the efficiency and accuracy of plant leaf disease identification, providing robust technical support for agricultural management.

6. Conclusion and future work

Existing methods for leaf disease recognition often design specific models for each plant, which is both time-consuming and labour-intensive. In addition, integrating multiple disease recognition models on a single device can lead to storage constraints. To address these problems, a novel plant leaf disease recognition model, termed PDNet, was proposed. The Adapter block is parameter-efficient, enabling PDNet to adapt to various plant disease recognition tasks with minimal parameter fine-tuning. The OPE ensures a reasonable patch division, allowing adjacent patches to overlap to some extent, thereby preserving critical disease features. In addition, the A-Softmax Loss, by introducing more stringent decision criteria, enhances the ability of the model to distinguish among similar leaf diseases.

This study represented a novel approach by employing the same network with nearly identical weights across various plant species. This

approach is expected to significantly advance leaf disease identification. The main limitation in this study is the high number of parameters in ViT itself. In the future, more attention will be given to minimizing the model's parameter count to achieve further improvements in performance.

CRedit authorship contribution statement

Xingshi Xu: Conceptualization, Formal analysis, Investigation, Methodology, Project administration, Software, Visualization, Writing – original draft, Writing – review & editing. **Yunfei Wang:** Data curation, Funding acquisition, Investigation, Writing – review & editing. **Yuying Shang:** Methodology, Writing – original draft, Writing – review & editing. **Zhixin Hua:** Data curation, Formal analysis, Investigation. **Zheng Wang:** Data curation, Formal analysis, Methodology. **Huaibo Song:** Conceptualization, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Validation, Writing – original draft, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by the National Key R&D Program of China (No. 2023YFD1301800), the National Natural Science Foundation of China (No.32272931), the Shaanxi Province Agricultural Key Core Technology Project (2023NYGG005), and the Shaanxi Provincial Technology Innovation Guidance Planned Program (2022QFY11-02). The authors would also like to thank the helpful comments and suggestions provided by all the authors cited in this article and the anonymous reviewers.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.engappai.2024.109466>.

Data availability

Data will be made available on request.

References

- Ashwini, C., Sellam, V., 2023. EOS-3D-DCNN: Ebola optimization search-based 3D-dense convolutional neural network for corn leaf disease prediction. *Neural Comput. Appl.* 35, 11125–11139.
- Chen, L., Zou, J.X., Yuan, Y., He, H.Y., 2023. Improved domain adaptive rice disease image recognition based on a novel attention mechanism. *Comput. Electron. Agric.* 208, 107806.
- Chen, S., Ge, C., Tong, Z., Wang, J., Song, Y., Wang, J.W., Luo, P., 2022. AdaptFormer: adapting vision transformers for scalable visual recognition. In: 2022 Conference on Neural Information Processing Systems. NeurIPS. MIT Press.
- Chen, Y.P., Wu, Q.F., 2023. Grape leaf disease identification with sparse data via generative adversarial networks and convolutional neural networks. *Precis. Agric.* 24 (1), 235–253.
- Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., Uszkoreit, J., Houlsby, N., 2020. An Image Is Worth 16x16 Words: Transformers for Image Recognition at Scale arXiv preprint arXiv:2010.11929.
- Ernest, M., Timnit, G., Andrea, F., Solomon, N., Jeremt, T., Chris, O., 2019. iCassava 2019 Fine-grained Visual Categorization Challenge arXiv preprint arXiv: 1908.02900.
- Feng, Q., Xu, P.F., Ma, D.X., Lan, G.Z., Wang, F.Y., Wang, D.W., Yun, Y.L., 2023. Online recognition of peanut leaf diseases based on the data balance algorithm and deep transfer learning. *Precis. Agric.* 24 (2), 560–586.
- Ferentinos, K.P., 2018. Deep learning models for plant disease detection and diagnosis. *Comput. Electron. Agric.* 145, 311–318.

- Gao, P., Geng, S., Zhang, R., Ma, T., Fang, R., Zhang, Y., Li, H., Qiao, Y., 2024. CLIP-Adapter: better vision-language models with feature adapters. In: 2023 Conference on International Journal of Computer Vision. IJCV. Springer.
- Guo, Y.F., Lan, Y.T., Chen, X.D., 2022. CST: convolutional Swin Transformer for detecting the degree and types of plant diseases. *Comput. Electron. Agric.* 202, 107–407.
- Hawi, G., 2021. Wheat leaf dataset. Mendeley Data V1. <https://doi.org/10.17632/wgd17666f17638n17636h.17631>.
- He, J., Chen, J., Liu, S., Kortylewski, A., Yang, C., Bai, Y., Wang, C., 2021. TransFG: A transformer architecture for fine-grained recognition. *arXiv preprint arXiv:2103.07976*.
- Houlsby, N., Giurgiu, A., Jastrzebski, S., Morrone, B., de Laroussilhe, Q., Gesmundo, A., Attariyan, M., Gelly, S., 2019. Parameter-efficient Transfer Learning for NLP *arXiv preprint arXiv:1902.00751*.
- Hughes, D.P., Salathe, M., 2015. An open access repository of images on plant health to enable the development of mobile disease diagnostics. *arXiv preprint arXiv:1511.08060*. <https://doi.org/10.48550/arXiv.1511.08060>.
- Jia, M., Tang, L., Chen, B., Cardie, C., Belongie, S., Hariharan, B., Lim, S., 2022. Visual prompt tuning. In: 2022 Conference on European Conference on Computer Vision. ECCV. Springer.
- Jiang, Z.C., Dong, Z.X., Jiang, W.P., Yang, Y.Z., 2021. Recognition of rice leaf diseases and wheat leaf diseases based on multi-task deep transfer learning. *Comput. Electron. Agric.* 186, 106184.
- Jin, H.B., Chu, X.Q., Qi, J.F., Zhang, X.X., Mu, W.S., 2023. CWAN: Self-supervised learning for deep grape disease image composition. *Eng. Appl. Artif. Intell.* 123, 106458.
- Kaur, P., Harnal, S., Gautam, V., Singh, M.P., Singh, S.P., 2022. An approach for characterization of infected area in tomato leaf disease based on deep learning and object detection technique. *Eng. Appl. Artif. Intell.* 115, 105210.
- Li, M.X., Zhou, G.X., Chen, A.B., Li, L.J., Hu, Y.H., 2023. Identification of tomato leaf diseases based on LMBRNNet. *Eng. Appl. Artif. Intell.* 123, 106195.
- Li, X.P., Chen, X.Y., Yang, J.L., Li, S.Q., 2022. Transformer helps identify kiwifruit diseases in complex natural environments. *Comput. Electron. Agric.* 200, 107258.
- Linderman, G.C., Steinerberger, S., 2017. Clustering with t-SNE. *arXiv preprint arXiv:1706.02582*. <https://doi.org/10.48550/arXiv.1706.02582>.
- Liu, W., Wen, Y., Yu, Z., Li, M., Raj, B., Song, L., 2017. SphereFace: Deep Hypersphere Embedding for Face Recognition *arXiv preprint arXiv:1704.08063*.
- Liu, Y., Ma, C., Tian, J., He, Z., Kira, Z., 2022. Polyhistor: parameter-efficient multi-task adaptation for dense vision tasks. In: 2022 Conference on Neural Information Processing Systems. NeurIPS. MIT Press.
- Long, M., OuYang, C., Liu, H., Fu, Q., 2018. Image recognition of *Camellia oleifera* diseases based on convolutional neural network & transfer learning. *Transactions of the CSAE* 34 (18), 194–201.
- Lu, X.Y., Yang, R., Zhou, J., Jiao, J., Liu, F., Liu, Y.F., Su, B.F., Gu, P.W., 2022. A hybrid model of ghost-convolution enlightened transformer for effective diagnosis of grape leaf disease and pest. *Journal of King Saud University-Computer and Information Sciences* 34 (5), 1755–1767.
- Luo, G., Huang, M., Zhou, Y., Sun, X., Jiang, G., Wang, Z., Ji, R., 2023. Towards efficient visual adaption via structural re-parameterization. *arXiv preprint arXiv:2302.08106*. <https://doi.org/10.48550/arXiv.2302.08106>.
- Medhi, E., Deb, N., 2022. PSFD-Musa: a dataset of banana plant, stem, fruit, leaf, and disease. *Data Brief* 43, 108427.
- Pan, S.Q., Qiao, J.F., Wang, R., Yu, H.L., Wang, C., Taylor, E.R., Pan, H.Y., 2022. Intelligent diagnosis of northern corn leaf blight with deep learning model. *J. Integr. Agric.* 21 (4), 1094–1105.
- Sachin, G., Ashish, K.T., 2024. Fruit and vegetable disease detection and classification: recent trends, challenges, and future opportunities. *Eng. Appl. Artif. Intell.* 133, 108260.
- Sathya, K., Rajalakshmi, M., 2022. RDA-CNN: enhanced super resolution method for rice plant disease classification. *Comput. Syst. Sci. Eng.* 42 (1), 33–47.
- Sethy, P.K., Barpanda, N.K., Rath, A.K., Behera, S.K., 2020. Deep feature based rice leaf disease identification using support vector machine. *Comput. Electron. Agric.* 175, 105527.
- Sladojevic, S., Arsenovic, M., Anderla, A., Culibrk, D., Stefanovic, D., 2016. Deep neural networks based recognition of plant diseases by leaf image classification. *Comput. Intell. Neurosci.* 2016, 3289801.
- Sudhesh, K.M., Sowmya, V., Kurian, P.S., Sikha, O.K., 2023. AI based rice leaf disease identification enhanced by Dynamic Mode Decomposition. *Eng. Appl. Artif. Intell.* 120, 105836.
- Sun, J., Tan, W., Mao, H., Wu, X., Chen, Y., Wang, L., 2017. Recognition of multiple plant leaf diseases based on improved convolutional neural network. *Trans. Chin. Soc. Agric. Eng.* 33 (19), 209–215.
- Sunil, C.K., Jaidhar, C.D., Nagamma, P., 2023. Tomato plant disease classification using Multilevel Feature Fusion with adaptive channel spatial and pixel attention mechanism. *Expert Syst. Appl.* 228, 120381.
- Tavakoli, H., Alirezazadeh, P., Hedayatipour, A., Nasib, A.H.B., Landwehr, N., 2021. Leaf image-based classification of some common bean cultivars using discriminative convolutional neural networks. *Comput. Electron. Agric.* 181, 105935.
- Thai, H.T., Le, K.H., Nguyen, N.L.T., 2023. FormerLeaf: an efficient vision transformer for cassava leaf disease detection. *Comput. Electron. Agric.* 204, 107518.
- Thakur, P.S., Khanna, P., Sheorey, T., Ojha, A., 2022. Trends in vision-based machine learning techniques for plant disease identification: a systematic review. *Expert Syst. Appl.* 208, 118117.
- Thangaraj, R., Anandamurugan, S., Kaliappan, V.K., 2021. Automated tomato leaf disease classification using transfer learning-based deep convolution neural network. *J. Plant Dis. Prot.* 128 (1), 73–86.
- Verma, S., Kumar, P., Singh, J.P., 2023. A meta-learning framework for recommending CNN models for plant disease identification tasks. *Comput. Electron. Agric.* 207, 107708.
- Vivek, S., Ashish, K.T., Himanshu, M., 2022. Technological revolutions in smart farming: current trends, challenges & future directions. *Comput. Electron. Agric.* 201, 107217.
- Vivek, S., Ashish, K.T., Himanshu, M., 2023a. CLDA-Net: a novel citrus leaf disease attention network for early identification of leaf diseases. In: 2023 15th International Conference on Computer and Automation Engineering. ICCAE, Sydney, Australia, pp. 178–182, 2023.
- Vivek, S., Ashish, K.T., Himanshu, M., 2023b. DLMC-Net: deeper lightweight multi-class classification model for plant leaf disease detection. *Ecol. Inform.* 75, 102025.
- Xu, J., Shao, M., Wang, Y., Han, W., 2020. Recognition of corn leaf spot and rust based on transfer learning with convolutional neural network. *Transactions of the CSAM* 51 (2), 230–236+253.
- Xu, X., Deng, H., Wang, Y., Zhang, S., Song, H., 2024a. Boosting cattle face recognition under uncontrolled scenes by embedding enhancement and optimization. *Appl. Soft Comput.* 164, 111951.
- Xu, X., Wang, Y., Hua, Z., Yang, G., Li, H., Song, H., 2023. Light-weight recognition network for dairy cows based on the fusion of YOLOv5s and channel pruning algorithm. *Transactions of the CSAE* 39 (15), 153–163.
- Xu, X., Wang, Y., Shang, Y., Yang, G., Hua, Z., Wang, Z., Song, H., 2024b. Few-shot cow identification via meta-learning. *Inform. Process. Agric.* 4 (2024), 1–11.
- Yang, B., Wang, Z.L., Guo, J.Y., Guo, L.L., Liang, Q.K., Zeng, Q., Zhao, R.Y., Wang, J.W., Li, C.H., 2023. Identifying plant disease and severity from leaves: a deep multitask learning framework using triple-branch Swin Transformer and deep supervision. *Comput. Electron. Agric.* 209, 107809.
- Zhang, G., Lv, Z., Liu, H., Liu, W., Long, C., Huang, C., 2023a. Model for identifying lotus leaf pests and diseases using improved DenseNet and transfer learning. *Trans. Chin. Soc. Agric. Eng.* 39 (8), 188–196.
- Zhang, X., Cen, C., Li, F., Liu, M., Mu, W., 2023b. CRFormer: cross-Resolution Transformer for segmentation of grape leaf diseases with context mining. *Expert Syst. Appl.* 229, 120324.
- Zhang, X.X., Xun, Y., Chen, Y.H., 2022. Automated identification of citrus diseases in orchards using deep learning. *Biosyst. Eng.* 223, 249–258.
- Zhou, C.J., Zhong, Y.J., Zhou, S.H., Song, J., Xiang, W.S., 2023. Rice leaf disease identification by residual-distilled transformer. *Eng. Appl. Artif. Intell.* 121, 106020.
- Zhou, M., 2019. Apple Foliage Diseases Recognition in Android System with Transfer Learning-Based. Univ. of NWFU master thesis.