# Algorithmic Approaches to Plant Disease Detection: A Comparative Analysis

*Jiaxu* Shen[1*]

[1]Beijing Academy, 100018 Beijing, China

**Abstract.** Every year, plant diseases cause considerable economic damage to agriculture, and the identification of plant diseases in the past relied on experts to determine the diseases through checking abnormal features of plants manually, which is time-consuming and labor-intensive. With the development of machine learning, models based on machine learning algorithms can now to identify diseases quickly using plant images. There are some algorithms used in identifying diseases, such as Convolutional Neural Networks (CNN), vision transformer (ViT), and K-means clustering. This study reviews the work already done by researchers. In turn, this work summarizes the advantages and limitations of some past studies and compares the popular algorithm CNN and other new algorithms 'performances through difference dimensions. By concluding the techniques used in different studies to enhance accuracy and ability of generalization thorough image preprocessing, selection of datasets, and use of new algorithms, and this work explores the potential and advantages of various algorithms to inform the development of improved models. It identifies key research gaps and offers recommendations for future studies.

## 1 Introduction

Food shortage remains a pressing global challenge with severe implications for human well-being and societal stability [1]. The Food and Agriculture Organization (FAO) reports over 800 million people worldwide suffer chronic undernourishment, while an estimated 1.3 billion spend less than $1 daily on food—highlighting economic constraints that worsen food insecurity. Plant diseases further exacerbate this crisis: they destroy vast amounts of global food production annually, deepening shortages and inflicting heavy losses on agricultural, forestry, and related economic sectors. A notable example is soybean rust, a destructive disease that has harmed farmers globally; research shows reducing its infection rate by 20% could yield growers an extra ~$11 million. This underscores the vital role of early plant disease detection and identification—timely recognition enables prompt interventions to mitigate losses and ease food security pressures.

Historical evidence shows plant disease damage is amplified by multiple factors. Lack of crop diversity leaves agricultural systems fragile: overreliance on a few crops lets single diseases devastate harvests. Pathogen-friendly climates ,such as high humidity, mild

* Corresponding author: shenjiaxu080115@gmail.com

temperatures, speed up harmful organisms' reproduction and spread, intensifying their impact. Many plant varieties also lack genes to resist common pathogens, leaving them defenceless and at risk of wipeout in severe outbreaks. For regions dependent entirely on one crop for livelihoods, even one or two diseases can cause catastrophic losses, and in extreme cases, trigger famine or hinder social development.

The globalization of agricultural trade complicates disease management further. When crops are transported to new cultivation areas, they lose natural resistance to local pathogens. In new environments, these plants cannot fend off locally evolved bacteria or native pathogens, leading to unseen diseases. This forces growers to confront unfamiliar threats, increasing their workload and control difficulty — making high-accuracy, practical identification algorithms an urgent priority.

Current plant disease control relies on two main approaches: biological and chemical methods [2,3]. Chemical control uses pesticides to suppress diseases but harms non-target plants, pollutes soil and water, and fosters pathogen resistance over time. Biological control — using beneficial organisms like insects or microorganisms — is eco-friendly, harmless to non-target plants, and as effective as chemical methods. However, its effectiveness varies by host plant genotype; a method working for one variety may fail for another. Thus, accurately distinguishing diseased from healthy plants is key to evaluating biological control success, further emphasizing the importance of reliable disease identification.

Among identification algorithms, Convolutional Neural Networks (CNNs) are widely adopted for their strong image processing ability—critical for identifying diseases from leaf or crop images. Traditional neural networks, by contrast, face limitations as layers increase, they fall into local optimization traps, suffer overfitting (poor performance on new data), and encounter gradient issues that hinder training [4]. CNNs address these via parameter sharing (reducing parameters and speeding convergence) and pooling operations (eliminating redundancy to prevent overfitting). Yet many CNN studies use limited datasets ,such as two datasets, raising doubts about their generalizability on unseen data.

Other algorithms show promise but have flaws. K-means clustering has low training costs and high accuracy but has only been tested on the simple "PlantVillage" dataset, failing to prove performance in complex real-world scenarios [5]. Self-supervised clustering, which integrates CNN advantages, avoids labor-intensive expert-labeled datasets (using unrefined data instead) and is fast with high accuracy. However, it has only been tested on one dataset, leaving its generalizability unvalidated.

This study will compare algorithms by accuracy, generalization, cost, and speed to guide their optimization. By linking performance to specific agricultural scenarios, such as small family farms, large commercial plantations, it bridges academic research and on-farm use, ensuring technical advances become actionable for growers. Ultimately, the research seeks to enhance the practicality of algorithm-based plant disease identification, enabling timely control, reducing crop losses, improving agricultural productivity, and supporting global food security.

## 2 Related work

### 2.1 CNN in plant disease identification

CNN consist of three core components: activation functions, optimizers, and loss functions. Activation functions simulate the transmission of information between different neurons; optimizers and loss functions work together to achieve the desired optimal learning outcomes [4]. CNNs reduce the amount of information required and minimize parameter counts. There are various types of CNNs, each with distinct advantages: Deformable convolution improves computers' ability to extract information from images, and depth wise convolution could reduce the cost when training machine learning models. Group convolutions have high accuracy than other algorithms without group convolution. Through these convolutional algorithms, people could recognize pictures, which could be utilized in discerning plant diseases.

In a research, researchers use two datasets to compares the ability of different models to distinguish diseases [6]. The researchers use a method called few-shot learning, which means the convolutional data are trained with two different data sources, one is unrelated with the target domain dataset, which is a major part of the data to train the model. The other part of the data is in the target domain dataset, but the amount of the data is very small. Thus, researchers could train a model with limited samples to get accurate results. In the experiment, at first, both the classifier and the feature extractor are trained. Then, only the classifier is trained with target domain dataset. The researchers use two classifiers, one is linear, and one is cosine-similarity based, which is called Baseline and Baseline++, respectively. An equation is used in linear classifier:

$$\tilde{Y} = \sigma(W_s^T f(X_i, \theta)) \tag{1}$$

In equation (1), is the weight matrix, d is the dimension of the feature vector and c is the number of classes.

In Baseline++ classifier, cosine similarity is calculated. And probability is obtained by normalizing the similarity vector. An equation is used to in Baseline++ classifier:

$$cs_{ij} = \frac{f(\mathrm{X}_i, \theta)^T w_{sj}}{\|f(\mathrm{X}_i, \theta)\| \|w_{sj}\|} \tag{2}$$

In the Baseline++ classifier, the weight matrix $W_s$ contains of c, d-dimensional weight vectors $[W_{s1}, W_{s2}, \cdots, W_{sc}]$. Every weight vector $W_{sj}$ could become as a prototype for one class. During the training process, calculation of the cosine similarity $CS_{ij}$ between the feature vector $Fi$ and the weight vector $W_{sj}$ is conducted as in equation (2).

Deep Adversarial Metric Learning (DAML) and Triplet Network are also included in the study. In the study, the researchers compare the ability of these models to identify plant diseases in two different settings, firstly, the models identify the diseases based on the plant type and disease name. Then, the models identify the diseases based on only common disease names, which is more a situation more common.

**Table 1.** Mean error of classification based on both crop and disease types (PlantVillage target domain dataset).

| **Approach** | **ResNet18/(shots)** | | | **ResNet25/(shots)** | | | **ResNet50/(shots)** | | |
|---|---|---|---|---|---|---|---|---|---|
| | 5 | 25 | 50 | 5 | 25 | 50 | 5 | 25 | 50 |
| Baseline | 17.78 | 5.58 | 4.24 | 14.6 | 4.38 | 2.52 | 14.42 | 3.98 | 2.48 |
| Baseline++ | 39.48 | 20.92 | 14.58 | 44.22 | 18.4 | 13.16 | 43.62 | 19.92 | 14.02 |
| TripletNetwork | 18.06 | 8.04 | 5.06 | 18.1 | 7.8 | 4.28 | 18.14 | 6.74 | 4.8 |
| DAML | 18.38 | 7.9 | 5.52 | 16.54 | 6.7 | 4.46 | 16.84 | 7.82 | 4.92 |

In Table 1, the models identified both plant types and disease types. Therefore, these models could identify the diseases of plants in fields directly. Baseline model has the lowest error: 14.42, 3.98, and 2.48.

 https://doi.org/10.1051/itmconf/20258001032

**Table 2.** Mean error of Classification based on both crop and disease types (Coffee Leaf target domain dataset).

| Approach | ResNet18/(shots) | | | ResNet25/(shots) | | | ResNet50/(shots) | | |
|---|---|---|---|---|---|---|---|---|---|
| | 5 | 25 | 50 | 5 | 25 | 50 | 5 | 25 | 50 |
| Baseline | 44.78 | 29.46 | 25.08 | 36.68 | 25.76 | 22.66 | 35.54 | 22.62 | 20.22 |
| Baseline++ | 54.96 | 43.84 | 38.38 | 56.7 | 38.5 | 33.32 | 60.12 | 38.32 | 32.7 |
| Triplet Network | 43.02 | 27.6 | 23.5 | 45.9 | 28.9 | 23.96 | 47.14 | 30.74 | 24.48 |
| DAML | 43 | 28.2 | 23.58 | 42.52 | 27.14 | 21.78 | 43.82 | 28.58 | 23.14 |

In Table 2, the models identify diseases through Coffee leaf. The lowest error also occurred in Baseline model: 35.54, 22.62, 20.22.

Based on different datasets to classify plant diseases, the baseline model has the lowest error for different shots compared to other models. In conclusion, under few-shot learning settings, the Baseline model outperformed all other models tested. The choose of models and training method could impact the accuracy a lot.

### 2.2 Key factors influencing the effectiveness of machine learning-based plant disease identification

In one study, the research determines How the use of different datasets, training strategies, and models impact the ability of generalization of models [7]. Models trained on the PlantVillage dataset achieved high accuracy and low overfitting under controlled conditions but exhibited a significant accuracy drop when applied to field conditions. PlantDoc datasets has some variations in the datasets, which result in low performance of models using the PlantDoc datasets. By comparing models using the data with background and without background, removing background could also increase the accuracy greatly. The study also fused multiple datasets to train a single model, finding that combining controlled-environment and field datasets substantially improved the model’ s overall accuracy. The research also combines a lot of datasets to train a single model and find that combining datasets with field condition and those without field conditions could increase the overall accuracy of the model.

## 3 Challenges & Application

### 3.1 Limitations of plant disease identification models

CNN is extensively used in the identification of plant diseases, but models based on ViT, which relies on self-attention mechanism, rather than traditional visual recognition paradigms, could captured global relationships of different patches of pictures [8]. The models based on ViT surpass traditional CNN models in identifying plant diseases. In an experiment, 46,409 pictures in both controlled conditions and field conditions were used to train pre-trained GoogLeNet CNN architecture through transfer learning, through removing the background, the accuracy is increased. However, the accuracy is still relatively low [9]. Thus, it is difficult to use CNN to identify a plant’ s various diseases through a picture of a leaf of the plant at once simply increasing dataset diversity. Moreover, some models of CNN are trained and tested in the same datasets, and the plants in pictures are placed in a same direction and position. Thus, the models fail to achieve high accuracy if they are trained and tested in different datasets in real field condition. Besides, using only finely processed pictures could also exacerbates the scarcity of available training datasets that could be used to train models.

### 3.2 Clustering-based models in plant disease identification

In one study, the researcher uses automatic K-means clustering to determine the plant diseases [5]. The study’ s goal is to analyze the effectiveness of background removal, image segmentation, feature extraction, and Principal Component Analysis (PCA). To remove the background, the study combined pixel clustering and edge detection. Image segmentation divides training images into multiple regions. K-means clustering first defines K-clusters, then classifies objects by calculating the squared distance between each object and cluster centroids, which includes several steps. K clustering also could select the diseases features from the healthy area of a plant, which could reduce the time to discerning the area of disease manually. Only some information is useful in determining the plant diseases. principle component analysis could convert a large dataset into a smaller one, which could reduce the training time and improve the ability of generalization. In feature selection, the important features to determine a disease of a plant could be determined, in turn the accuracy could be increased. The results of the experiment showed that the removal of the background and division of images could increase the accuracy in a very short time. The K-means clustering model trained on background-free images achieved higher accuracy than other groups. The performance of K-means clustering is also compared with the performance of CNN and GoogLeNet, CNN and GoogleNet are time consuming and have lower accuracy. Dimensionality reduction of the dataset to reduce the size of the dataset is proved to be useful in reduce the time to train models, and feature selection also result in 60% reduction in time of diagnosis. Thus, by using different training strategies and k-mean clustering, the efficiency, accuracy could be improved.

### 3.3 Application of ViT in plant disease classification

In the first study, researchers employed the inverted residual block technique, which increases channel dimension, facilitates the capture of long-range pixel dependencies, reduces computational complexity, and enhances information inference efficiency [8]. Mobile ViT block is used to avoid the loss of information through splicing original feature maps, while reduce the computational costs. The identification process also includes Vision transformer encoder. Firstly, the pictures are divided into batches, which is then flattened. Then, the SoftMax function is used:

$$self-attention = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V \tag{3}$$

In equation (3), the attention scores could be obtained, which could be used to calculate the feature information.

CBAM block is used in the study to preprocess the images before convolution kernel application, and the channels are compressed and then expanded to original number of channels, and the weight of the channels is calculated through the equation (4):

$$M_c(F) = \sigma(MLP(AvgPool(F)) + MLP(MaxPool(F))) = \sigma(W_1(W_0(F^c_{avg})) + W_1(W_0(F^c_{\max}))) \tag{4}$$

Equation (4) illustrates the module’s weight assignment process. σ represent that the function Sigmoid is used as the activation function, $W_1$ ∈ R^(C/r × C), and $W_0$ ∈ R^(C/r × C).the shared weights for the two inputs of the max pooling layer and the average pooling layer are $W_1$ and $W_0$.

The study compared their model's accuracy with numerous other CNN-based and ViT-based models. Moreover, the model was tested with images of different parts of different crops, including wheat, rice, and coffee, which enables the researchers to compare the models' accuracy in different background and field environment. The researcher's model has the highest accuracy in determining the diseases compared to models with similar parameters and had significantly fewer parameters than heavyweight network. Thus, the study's model enables efficient, cost-effective, and accurate plant disease identification under both laboratory and real-world natural conditions—an essential capability for modern plant disease control.

In the second study, the researchers compare models based on ViT and CNN to identify plant diseases [10]. The images of the datasets used in the study is processed first into vectors by flattening fixed sized blocks, which serve as the inputs of the model. Then, the positional encodings are used to determine the position of each patch, and the transformer encoder is introduced, which is the essential step of the model, which help to create feature vectors, and the optimizer is used to train the model, changing the weights used in embedding and transformer layers. In the study, the researchers find that the accuracy of the ViT model surpass the accuracy of the inception V3 model using CNN algorithm. When the models are tested by a dataset that is not included in training process, the ViT model also get higher accuracy and lower loss rate than the inception V3 model.

# 4 Suggestion

## 4.1 Selection of the datasets

Most experiments identifying plants diseases used a few commonly used datasets, like PlantVillage, which is a dataset in experimental environment: in the datasets, it contains various plants species with different diseases, and each photograph only shows one specific part of the plants: a piece of leaf. Thus, models could be trained effectively without many disturbances. Some datasets, however, contain several parts of the plants, some a lot of leaves in a single plant, which could simulate pictures in fields, and images and these plants are pictured in different weather conditions, angles, and light intensity, which is more like images taken in field conditions. Thus, models trained by these datasets could exhibit higher identification accuracy when they are used. A study shows that when a model is only trained by datasets in experimental environment, achieve high accuracy when it is tested with the same datasets but get lower accuracy when it is tested with other datasets. However, models trained by both datasets in experimental environment and field conditions have higher accuracy when it is tested by the same datasets and different datasets. Moreover, because factors like climate and differences in plant species in different areas, using datasets that have data from various regions of the world could enhance the accuracy and generalization ability. Thus, if future studies could train and test models with experimental and field conditional datasets from different places, the most suitable algorithms could be found, and the overall accuracy of models could increase.

## 4.2 Preprocessing of datasets

Most datasets include pictures that include the plant itself and the background, and the ground is not account for the identification of plant diseases, which means that background is irrelevant data that could disturb models' determination of diseases. Furthermore, the background of the datasets should be removed, to prevent models from being affected by unrelated background features. In a study, researchers determined several models' accuracy in identification of plant diseases and found that models trained on background-removed datasets achieved higher accuracy. Feature selection is another technique that emphasizes key features that are important in disease identification, like the features of plants that are typical symptoms of some disease and ignore the irrelevant features. Another challenge, thereby improving models to increase their accuracy. Moreover, in a study, plant disease regions are marked by experts in plant disease identification before training the models, which also could enhance the accuracy. Thus, preprocessing the datasets used for training through removing backgrounds and feature selection could improve the accuracy of models. Moreover, a study applied these pictures with different angles, flip them randomly to alleviate overfitting and reduce training time.

### 4.3 Optimization of CNN-based plant disease identification

There are many studies using CNN to identify plant diseases, and traditional CNN in turn has been improved a lot. CNN could analyze and extract the features of plant diseases through limited available data, which enables them to detect some rare and new diseases with few recorded samples. The problem that some diseases have plenty of samples and some have little also could be solved through CNN model through FL-Efficient Net network in one study. However, some of these studies only use one or two datasets to train and test their models, and these datasets are both images collected in laboratory environment. This leads to when the models are used in fields to detect plant diseases, the accuracy of these models might drop. Thus, in the future studies of CNN in plant disease identification, field-collected images into their training and test datasets should also be used to train or test the models. Moreover, one study indicates that if the pictures that contain several parts of plants are used to test or train models, CNN models' performance could be negatively influenced. When the CNN models are applied, users might take photos that contain several parts of plants that have diseases. Thus, In the future study, new methods should be developed to improve the ability of CNN models in identifying plants diseases with pictures containing multiple parts of the plants. Moreover, sometimes plants have several diseases at the same time, but models with CNN in one study only determine the dominant disease. models using CNN could develop techniques to discern several diseases of a single plant.

### 4.4 Enhancement of K-means clustering for plant disease identification

The model based on K-means clustering used automatic clustering to identify the disease area of plant leaves, which could save time and resources to diagnose these areas by experts manually, and the number of regions of interest (ROIs) of the K-means clustering algorithm could be determined automatically. CNN is a commonly used algorithm in training models to identify plant diseases. However, in the study using the K-means clustering model, the accuracy and time used in identifying plant diseases of the model using K-means clustering algorithm is higher of models that use CNN. However, the study only involves one dataset, and the research only select three types of diseases of leaves. This means that the study trains and tests the model with the same datasets, which could not guarantee that the model could have the same accuracy when it is tested by data selected from different conditions. Moreover, the datasets used by the study only contain pictures took in labs. Thus, the model might face difficulties when it is used to identify some diseases of plants in field conditions, which have a lot of factors, like light intensity, weather condition, and angle of lens, that impact the quality of pictures. Therefore, in the future studies, models based on K-means clustering should improve through be trained and tested by multiple types of datasets, and applications based on K-means clustering to identify plant diseases could be developed to use in production of crops. Accordingly, the ability of generalization of K-means clustering models could be improved.

### 4.5 Advancement of ViT-based plant disease identification algorithms

In a study, researchers compare the accuracy of CNN, and the algorithm based on ViT in identification of plant diseases, and found that the model based on ViT perform best, and the researchers develop a application as a viable method to use this model in reality. But the model based on ViT in the study only discern diseases of tomatoes and found that the model identified some diseased with greater accuracy than others. Thus, future studies could investigate the reasons behind it and develop new models to improve ViT-based models when identifying multiple diseases of various species.

## 5 Conclusion

In the study, models in identifying plant diseases based on CNN, ViT, and K-means clustering is analyzed. These utilized various techniques to increase the accuracy of the models, including selection of datasets, processing of datasets, and improvement of algorithms. Through training and testing the models with both datasets in experimental conditions and field conditions, the ability of generalization could be increased. Through extraction of features, augmentation, background removing, and other techniques, the accuracy of models could increase, because the noise of pictures, which is irrelevant to the diagnosis of plants, is handled and ignored by models when considering diseases of plants. Specifically, study develop a CNN based model that could identify plant diseases with only limited data from testing datasets. Through studying the datasets that is irrelevant to testing, the model could also achieve high accuracy. The model based on ViT use multiple techniques to extract the features from the datasets and achieve higher accuracy than the CNN based model used in the study. The model based on K-means clustering in a study could acquire the ROI number of the algorithm automatically, instead of calculating the numbers manually by professionals. Both studies have some deficiencies, the study based on CNN could only determine the major disease of a plants and analyze one part of the plant once. The study based on ViT and K-means clustering use limited datasets to examine their models' accuracy. Therefore, through improving these models and explore potential suitable algorithms to identify plant diseases, the economic loss result from plant disease could be reduced because of timely crop disease treatment of crops through effective and convenient machine learning models to identify diseases.

## Reference

1. S. Peter, Plant Disease: A Threat to Global Food Security. Annu. Rev. Phytopathol. **43**, 83-116 (2005) https://doi:10.1146/annurev.phyto.43.113004.133839
2. T. Hirooka, H. Ishii, Chemical control of plant diseases. J. Gen. Plant Pathol. **79**, 390–401 (2013) https://doi:10.1007/s10327-013-0470-6
3. P.A. O'Brien, Biological control of plant diseases. Australas. Plant Pathol. **46**, 293–304 (2017) https://doi:10.1007/s13313-017-0481-4
4. Z. Li, F. Liu, W. Yang, S. Peng, J. Zhou, A Survey of Convolutional Neural Networks: Analysis, Applications, and Prospects. IEEE Trans. Neural Netw. Learn. Syst. **33**, 6999-7019 (2022) https://doi:10.1109/TNNLS.2021.3084827

5. S.M. Javidan, A. Banakar, K.A. Vakilian, Y. Ampatzidis, Diagnosis of grape leaf diseases using automatic K-means clustering and machine learning. Smart Agric. Technol. **3**, 100081 (2023) https://doi:10.1016/j.atech.2022.100081
6. A. Afifi, A. Alhumam, A. Abdelwahab, Convolutional Neural Network for Automatic Identification of Plant Diseases with Limited Data. Plants. **10**, 28 (2021) https://doi:10.3390/plants10010028
7. A. Ahmad, A.E. Gamal, D. Saraswat, Toward Generalization of Deep Learning-Based Plant Disease Identification Under Controlled and Field Conditions. IEEE Access. **11**, 9042-9057 (2023) https://doi:10.1109/ACCESS.2023.3240100
8. U. Barman, P. Sarma, M. Rahman, V. Deka, S. Lahkar, V. Sharma, M.J. Saikia, ViT-SmartAgri: Vision Transformer and Smartphone-Based Plant Disease Detection for Smart Agriculture. Agronomy. **14**, 327 (2024) https://doi:10.3390/agronomy14020327
9. S.B. Jadhav, V.R. Udupi, S.B. Patil, Identification of plant diseases using convolutional neural networks. Int. J. Inf. Technol. **13**, 2461 – 2470 (2021) https://doi:10.1007/s41870-020-00437-5
10. G. Li, Y. Wang, Q. Zhao, P. Yuan, B. Chang, PMVT: a lightweight vision transformer for plant disease identification on mobile devices. Front. Plant Sci. **14**, 1256773 (2023) https://doi:10.3389/fpls.2023.1256773