# Revolutionizing Plant Disease Diagnosis: An Overview of Automated Techniques for Leaf Disease Classification and Prediction System

[1]B.Dhivya, [2]E. Jean Edward Clement, [3]S.Sathiya, [4]Dr. H. Niroshini Infantia
[1]Department of AI&DS, Karpaga Vinayaga College of Engineering and Technology, Chengalpattu, Tamilnadu.
[4]Department of CSE, Assistant Professor, Roever Engineering College, Perambalur, Tamilnadu
[4]Department of CSE, Assistant Professor, Arasu Engineering College, Kumbakonam, Tamilnadu
[4]Associate Professor, St.Joseph's Institute of Technology, OMR, Chennai, Tamilnadu
[1]dhivyabaskar91@gmail.com, [2]jeanecm@gmail.com, [3]sathiya2022apc@gmail.com
[4]niroshiniinfantiah@gmail.com

**Abstract—** India is a country that is considered to be one of the developing nations in the world, and its economy is heavily reliant on agricultural production. The identification of crop diseases continues to be a difficult undertaking, despite the fact that crop diseases pose a significant risk to food security. Plant cultivation is negatively impacted by a number of diseases, including gray leaf spot, common rust, healthy leaf, and northern leaf blight, among others. As a consequence of this, the plant's leaves, roots, and stems are all affected by these diseases, which lead to a significant reduction in output. It is a difficult effort to classify diseases under human supervision because there is such a vast variety of diseases, and this result in a high rate of mistake.

Furthermore, in order to overcome these restrictions, we need to design an automated system for identifying and categorizing the diseases that affect the leaves. We have presented an overview of the plant illnesses, the many types of plant diseases, and the approaches that may be used to treat plant leaf diseases in this technical review paper. Furthermore, we have included a literature review that is already in existence, as well as a concise description of the technologies that are currently and recently deployed for the purpose of diagnosing and categorizing plant leaf diseases by making use of some benchmark datasets that are publically available.

**Keywords:** ***plant leaf, disease prediction, computer vision, predictions, machine learning and deep learning.***

## 1. Introduction

In the majority of developing countries, agriculture is the primary work that people do. Additionally, in order for humans to survive, food is one of the fundamental requirements, along with air, water, and clothing supplies. The government and the scientific community are focusing their attention on methods to improve agricultural yield as a result. If crops are not destroyed in fields by plant diseases or pests, then there is the potential for a rise in both the quality and quantity of agricultural produce. It has the potential to assist in a variety of ways, including enhancing the profits of farmers by increasing the quantity of crops that are of higher quality and that are sold in greater quantities, and supplying food to those who are less fortunate, given that there is a greater quantity of food in a nation [1-2].

As a result of contaminated crops, it also helps to prevent diseases that are related to food that can affect both humans and animals. It is therefore of the utmost importance to identify plant diseases because doing so could assist in the spraying of appropriate fungicides to safeguard the crops. Farmers have been able to improve their ability to cope with this issue by attending a number of plant clinics and making use of fertilizers and fungicides [3–5]. The procedure of identifying the illness that affects their crops is typically entrusted to plant scientists by farmers, and it is a laborious and time-consuming endeavor. An further point to consider is that the viewpoint of a plant scientist may vary from one individual to another. Agriculture scientists, after conducting a large number of experiments, came to the conclusion that the leaves of a plant are the most obvious sign of a disease that is hurting the plant. Fungal, bacterial, and viral diseases are the three categories that can be described as plant diseases [6–8]. Fungal infections account for approximately 85 percent of all diseases, and they are responsible for the visual impacts that are seen in plant leaves.

The diseases known as leaf rust and powdery mildew are examples of diseases that are caused by fungus. The yellow halo that surrounds a leaf spot can be caused by bacterial infections. Mosaic leaf patterns or yellowed leaves can also be caused by viral illnesses similarly. In spite of the fact that plant illnesses can manifest themselves visually in any part of the plant, including the leaves, stems, or fruits, the most common way for determining which disease is affecting crops is to observe the changes in the color or form of the leaves [9].

A web-based software solution that can be offered to farmers to know plant disease simply by taking images from high-intensity mobile cameras is a pressing requirement that has arisen as a result of the introduction of the internet into people's lives in the form of smartphones. It is also possible to design software in the form of an application, which can be conveniently utilized by the agricultural community [10]. Mobile applications have also been developed. Image processing [11–13] and the most recent Neural Network research findings have the potential to significantly enhance plant development and protection techniques. This is made possible by more powerful computing power. The remainder of the paper is organized based on the following structure: A summary of the various types of plant leaf diseases that have been observed over the years is presented in Section 2. A discussion of the many datasets that are available to the public is included in Section 3, along with information regarding the datasets themselves. In Section 4, we will discuss the present research that pertains to the many methods that may be used to diagnose the plant leaf disease. In Section 5, we will discuss the conclusions that can be drawn from this study as well as the potential future applications of this research.

## 2. Literature Survey

With regard to the detection and classification of plant diseases, there are two artificial intelligence (AI) methodologies that are commonly used, as demonstrated in Figure 1. In the realm of artificial intelligence, the most prevalent methodologies are Neural Networks, which include Deep Convolutional Neural Networks (Deep CNN) [6, 20–22], as well as Machine Learning techniques like Logistic Regressions, Decision Trees, Support Vector Machines (SVM) [5, 23], k-Nearest Neighbors (k-NN), Naïve Bayes [24], and other examples.

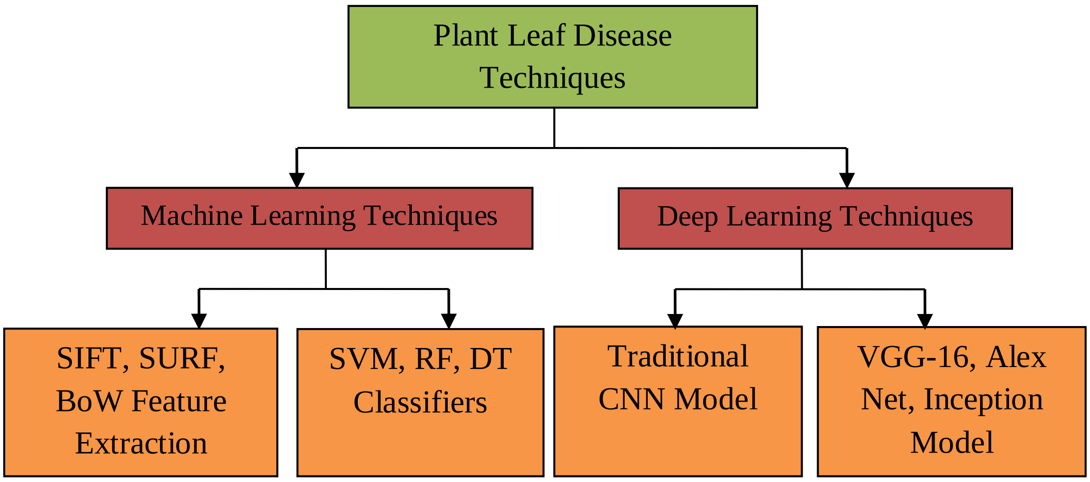


**Fig. 1** Plant leaf disease classification techniques

### 2.1. Machine Learning based Plant Disease Classification

Machine learning is one of the emerging methods that can be used to classify diseases that affect plant leaves. The procedures involved in the classification of plant leaf diseases using machine learning are graphically depicted in Figure 2. In the plant leaf disease classification method that is based on machine learning, there are three phases that need to be followed.

- Collecting the Dataset
- Pre- Processing the dataset
- Hand-craft feature extraction techniques
- Traditional classification
- Prediction result

Using images of the leaves, for instance, Rumpf et al. [14] demonstrated the application of support vector machines (SVMs) for the detection of three different illnesses in sugar beetroot. Mokhtar et al. [15] also employed support vector machines (SVM) to categorize two different types of tomato plant illnesses, and they ended up with a 92% accuracy rate. photos of leaves can also be classified in [16], where Pantazi et al. attempted to categorize photos of vines into four distinct categories. The colorful image is initially divided into a 32-bin histogram using a technique known as Local Binary Pattern (LBP). For each of the four different diseases, one class of support vector machines (SVM) were trained. A recognition rate of 97%, 95%, and 93% was found for sick leaves when tested on 100 photos of each class, while the recognition rate for healthy leaves was 100%. For the purpose of identifying powdery mildew illness in tomato plants, Prince et al. [17] have demonstrated that the SVM classifier with linear kernel provides an accuracy of more than 90 percent.

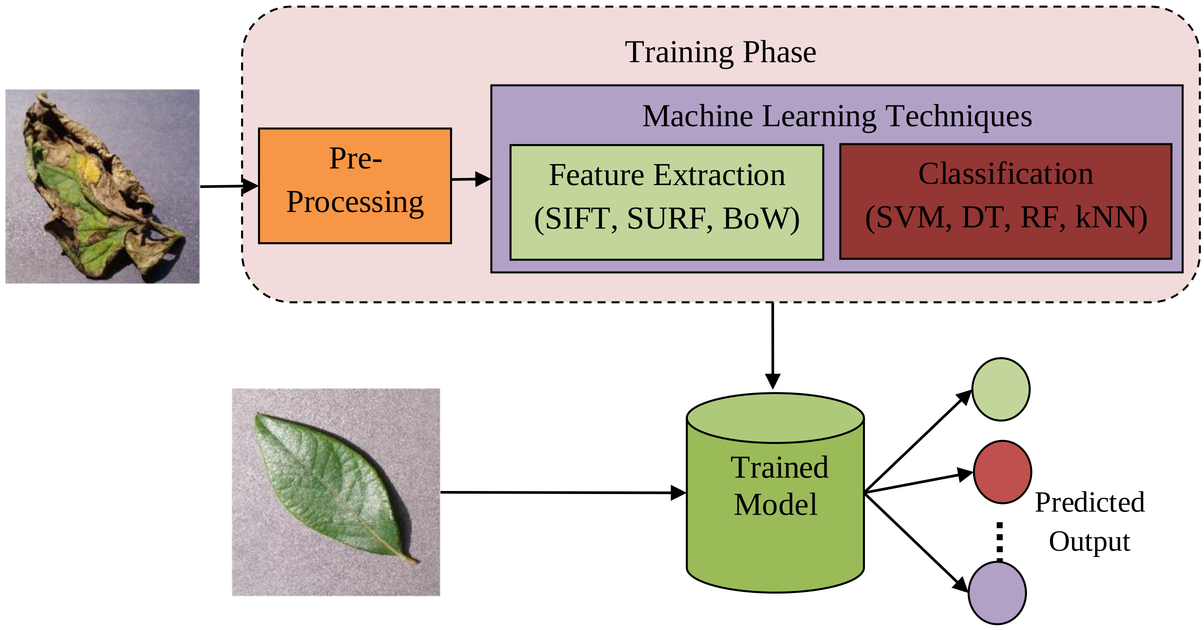


**Fig. 2** Sequence of machine learning based plant leaf disease recognition system

For the purpose of classifying three classes of rice plant diseases and one healthy class, Shrivastav and Pradhan [18] utilized seven different classifiers. These classifiers included

SVM, Discriminant Classifier, k-NN, Naive Bayes, Decision Trees, Random Forest, and Logistic Regression. As stated by the authors, the SVM classifier was able to reach a maximum accuracy of 94.65%. In the article [19], Yang and Guo provide a comprehensive review of various machine learning techniques, including the Naïve Bayes classifier, Support Vector Machine (SVM), K-Means clustering, Artificial Neural Networks (ANN), Decision Trees, and Random Forests. These techniques are put to the test in order to identify plant diseases based on leaf images. In order to identify genes that are involved in interactions between plants and pathogens (bacteria or viruses that cause disease), the scientists make use of machine learning methodology. For the purpose of identifying Huanglongbing in citrus orchards, the authors mentioned that Sankaran et al. [20] demonstrated the highest accuracy of 95% by employing Quadratic Discriminant Analysis (QDA).

**Table 1 Literature information about ML based Plant Leaf Disease System**

| S. No. | Authors | Leaf Disease | Model | Accuracy |
|---|---|---|---|---|
| 1. | Rumpf et al. [14] | Sugar beet | SVM | 89.5% |
| 2. | Mokhtar et al. [15] | Tomato leaf | SVM | 92% |
| 3. | Pantazi et al. [16] | Vine image | LBP and SVM | 93.15% |
| 4. | Prince et al. [17] | Tomato plant leaf | SVM with Linear kernel | 90% |
| 5. | Shrivastav and Pradhan [18] | Rice plant | SVM, k-NN, NB, DT, RF and LR | 94.65% |
| 6. | Yang & Guo [19] | Tomato plant leaf | SVM, RF, ANN, K-means | 85.2% |

### 2.2. Deep Learning based Plant Disease Classification

Recently, a number of researchers have shown that Deep Learning is a more effective method for achieving high accuracy in the identification of plant diseases. The majority of the time, researchers have reported positive outcomes in disease categorization when they applied transfer learning to models that had previously been trained in other domains. Using the PlantVillage dataset as an example, Mohanty et al. [21] have utilized well-known CNN models such as AlexNet and GoogLeNet in order to make predictions regarding the classifications of plants and diseases that are present in the collection. On a gray background, the PlantVillage dataset contains leaf photos of fourteen different crops, with each crop being classed into two to ten different disease categories. The usage of pre-trained models Alexnet and VGG (Visual Geometry Group) has been demonstrated by Ferentinos [22] for the classification of a plant as well as disease from 25 different plants and 58 separate classes of crop and disease combined. This classification was accomplished by using the models.

A plant leaf disease identification model that is based on a deep convolutional neural network (Deep CNN) was proposed by Geetharamani et al. [23]. This model obtains an average accuracy of 96.46%.The method shown that hybrid strategies are superior to any other deep learning techniques that are currently available. An accuracy of 99.35% was achieved by Sharada et al. [24] by the utilization of a public dataset consisting of 54306 photos of diseased and healthy plant leaves. These images were captured under controlled settings and were used to train a deep neural network.It is possible to minimize the size of the dataset by removing the repeated number of photos, which came to 41000.

The same thing happened with Deng R et al. [25], who produced automatic plant recognition by utilizing deep learning techniques. Through the use of deep learning models that are ensembles to locate beat modules, the proposed models were able to reach an accuracy of 91%. These models took into consideration the rice plant and were able to detect six different types of diseases. A big dataset of 33,026 photos of six different forms of rice illnesses was used to construct the approach. These diseases include brown spot, leaf blast, false smut, neck blast, sheath blight, and bacterial stripe disease. The method was produced using deep learning by using the dataset. Integrated submodels were incorporated into the Ensemble Model, which served as the method's central component. Last but not least, the Ensemble Model was confirmed by employing a distinct collection of photos.

A novel deep learning architecture known as Inc-VGGnet is presented by Junde et al. [26]. This architecture achieves a validation accuracy of less than 91.83% when applied to the public dataset. They came up with a model that is a combination of the pre-trained MobileNet platform and the SE block. In order to generate the best possible model, the transfer learning process was carried out twice. The identification of maize leaf disease from nine different varieties of maize leaves was accomplished by the use of transfer learning in Zhang et al. [27]. The authors report an accuracy of 98.9% for the GoogleLeNet dataset and 98.8% for the Cifar10 dataset when they use max-max-ave pooling in three CNN hidden layers simultaneously. Rangarajan et al. [28] have also documented the utilization of transfer learning. They have utilized AlexNet and VGG16 in order to train and identify photos of tomato plant diseases that were taken from the PlantVillage dataset. On a total of 373 test photos belonging to each category, they demonstrated that AlexNet achieves an accuracy of around 97.29%, whereas VGG16 achieves an accuracy of approximately 97.49%.

**Table 2 Literature information about DL based Plant Leaf Disease System**

| S. No. | Authors | Leaf Disease | Model | Accuracy |
|---|---|---|---|---|
| 1. | Mohanty et al. [21] | Plant Village Dataset | Alex Net and Google Net | 94.5% |
| 2. | Ferentinos [22] | Plant Village Dataset | Alex Net and VGG-16 Net | 95.1% |
| 3. | Geetharamani et al. [23] | Plant Village Dataset | Deep CNN Model | 96.46% |
| 4. | Sharada et al. [24] | Plant Village Dataset | Inception Net | 99.35% |
| 5. | Deng R et al. [25] | Rice plant leaf | Ensemble Model | 91% |
| 6. | Junde et al. [26] | Maize leaf plant | VGG-16 Net Model | 98.9% |

In terms of several characteristics, the deep learning models that were discussed earlier, such as Alex Net, DenseNet-121, VGG-16 Net, and ResNet-50, as well as ensemble models, have a number of limitations. The pre-trained model and the Ensemble Model have a number of parameters, which can slow down the identification process and require more time to train.

## 3. Recent Trends in Deep Learning Techniques

In the field of machine learning, deep learning refers to a technique that is utilized to construct a model that is capable of learning from many types of data, including photos, videos, audio, and text, and carrying out tasks such as classification and detection. A number of different scene classification models, such as Convolutional Neural Network (CNN) [29] and Deep CNN models [30], are utilized in the process of scene classification through the utilization of scene classification. The Convolutional Neural Network (CNN), transfer learning (VGG-16 and VGG-19) [31-33], Recurrent Neural Network (RNN), Auto Encoder, and Generative Adversarial Network are several types of techniques that are based on deep learning. Models that use CNN and transfer learning are frequently utilized for the purpose of picture classification. On a large scale, CNNs are utilized for the purpose of image segmentation and classification.

In spite of the fact that CNNs were initially developed in 1989, the remarkable success that they had in the ImageNet Competition in 2012 brought them even more attention. Increasing the number of layers, neurons with millions of weights, and connections between distinct neurons all contribute to an increase in the computational complexity of CNN architecture. Figure 8 depicts the fundamental block diagram of CNN, which consists of convolutional, pooling, activation function, and fully linked layers. Each of these layers serves a distinct function in the CNN.

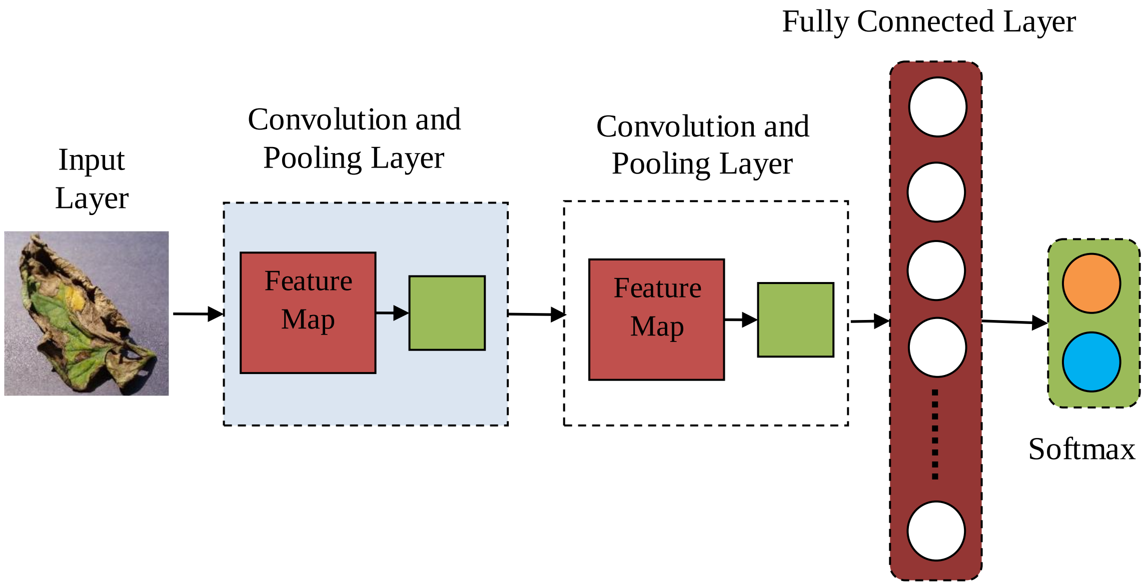


**Fig. 3** Working flow of Convolutional Neural Network for disease recognition system

The feature maps are generated by the convolutional layer through the process of convolving the input images across the kernel. In the process of transferring the value to the subsequent layer, the results from the convolutional layers that came before it are downsampled by using the maximum or average of the neighborhood that has been defined in the pooling layer respectively. For the purpose of providing a forecast of the input data, the loss function is employed in conjunction with the other layers of the CNN model output.

In conclusion, the parameters of the network are established by minimizing the loss function between the prediction labels and the ground truth labels while simultaneously maintaining the regularization requirements. Additionally, backpropagation is utilized in order to modify the weights of the network throughout each iteration until convergence is achieved. For instance, R. Sangeetha and M. Mary Shanthi created a deep learning technique that was based on a CNN model for the purpose of classifying plant disease-affected leaves as opposed to healthy leaves [34]. In a similar manner, Deepalakshmi et al. [35] developed an automated system that would extract features from input photos using the CNN algorithm in order to distinguish between damaged and healthy leaves of different plants.

## 4. Conclusion

During the development of this technical review study, we made an effort to examine the several methods that can be utilized to employ image processing techniques for the aim of identifying diseases that affect plant leaves. Over the past few years, there has been a significant amount of progress achieved in the field of automated or computer vision for the purpose of disease diagnosis and classification; yet, there is still room for further advancement. Unfortunately, there is no method that is completely effective in identifying all illnesses that can affect plant leaves.

We discovered that one of the standard methods is a disease identification system that is based on machine learning. Nevertheless, the machine learning approaches are challenging to manage when applied to huge datasets. The illness detection system that is based on deep learning is an emerging technology for managing enormous datasets. Additionally, the system will provide more accurate results than the ones that are produced by typical machine learning techniques.

## References


1. Kamilaris, A & Prenafeta Boldú, "Deep Learning in Agricultural: A Survey, Computers and Electronics in Agriculture, Vol. 147, pp. 70-90, 2018.
2. Mahlein, AK, Oerke, EC, Steiner, U & Dehne, Recent Advances in sensing plant diseases for precise crop production, Eur. Journal Plant Pathol, Vol. 133, pp. 197-209, 2012.
3. Kaur, S.; Pandey, S.; Goel, S. Plants disease identification and classification through leaf images: A survey. Arch. Comput. Methods Eng., Vol. 26, pp. 507–530, 2019.
4. Hamuda E., Glavin M., and Jones E., A Survey of Image Processing Techniques for Plant Extraction and Segmentation in the Field, Computers and Electronics in Agriculture, Vol. 125(1), pp. 184–199, 2016.
5. Barbedo, J.G.A. A review on the main challenges in automatic plant disease identification based on visible range images. Biosyst. Eng., Vol. 14(4), pp. 52–60, 2016.
6. Chouhan S., Singh U., and Jain S., Applications of Computer Vision in Plant Pathology A Survey, Archives of Computational Methods in Engineering, Vol. 27(2), pp. 611–632, 2017.

7. El Houby E.M.F., “A survey on applying machine learning techniques for management of diseases”, Journal of Application in Biomedical, pp.165–74, 2018.
8. Barbedo J., An Automatic Method To Detect and Measure Leaf Disease Symptoms Using Digital Image Processing, The American Phytopathoogical Society, Vol. 98(12), pp. 1709–1716, 2014.
9. A. A. Bharate and M. Shirdhonkar, “A review on plant disease detection using image processing,” in 2017 International Conference on Intelligent Sustainable Systems (ICISS). IEEE, 2017, pp. 103–109.
10. U. Mokhtar, M. A. Ali, A. E. Hassanien, and H. Hefny, “Identifying two of tomatoes leaf viruses using support vector machine,” in Information Systems Design and Intelligent Applications. Springer, 2015, pp. 771–782.
11. E. Hamuda, M. Glavin, and E. Jones, “A survey of image processing techniques for plant extraction and segmentation in the field,” Computers and Electronics in Agriculture, vol. 125, pp. 184–199, 2016.
12. J. I. Arribas, G. V. Sánchez-Ferrero, G. Ruiz-Ruiz, and J. Gómez-Gil, “Leaf classification in sunflower crops by computer vision and neural networks,” Computers and Electronics in Agriculture, vol. 78, no. 1, pp. 9–18, 2011.
13. Oo Y. and Htun N., (2018a), Plant Leaf Disease Detection and Classification Using Image Processing, International Journal of Research and Engineering, 5(9), 516–523.
14. T. Rumpf, A.-K. Mahlein, U. Steiner, E.-C. Oerke, H.-W. Dehne, and L. Plümer, “Early detection and classification of plant diseases with support vector machines based on hyperspectral reflectance,” Computers and Electronics in Agriculture, Vol. 74, no. 1, pp. 91–99, 2010.
15. U. Mokhtar, M. A. Ali, A. E. Hassanien, and H. Hefny, “Identifying two of tomatoes leaf viruses using support vector machine,” in Information Systems Design and Intelligent Applications. Springer, 2015, pp. 771–782.
16. X. E. Pantazi, D. Moshou, A. A. Tamouridou, and S. Kasderidis, “Leaf disease recognition in vine plants based on local binary patterns and one class support vector machines,” in IFIP International Conference on Artificial Intelligence Applications and Innovations. Springer, 2016, pp. 319–327.
17. G. Prince, J. P. Clarkson, and N. M. Rajpoot, “Automatic detection of diseased tomato plants using thermal and stereo visible light images,” PloS One, Vol. 10, No. 4, p. e0123262, 2015.

18. S. Shrivastava, S. K. Singh, and D. S. Hooda, “Color sensing and image processing based automatic soybean plant foliar disease severity detection and estimation,” Multimedia Tools and Applications, Vol. 74, No. 24, pp. 11 467–11 484, 2015.
19. X. Yang and T. Guo, “Machine learning in plant disease research,” European Journal of BioMedical Research, vol. 3, no. 1, pp. 6–9, 2017.
20. S. Sankaran, A. Mishra, J. M. Maja, and R. Ehsani, “Visible-near infrared spectroscopy for detection of huanglongbing in citrus orchards,” Computers and Electronics in Agriculture, vol. 77, no. 2, pp. 127–134, 2015.
21. S. P. Mohanty, D. P. Hughes, and M. Salathé, “Using deep learning for image-based plant disease detection,” Frontiers in Plant Science, vol. 7, p. 1419, 2016.
22. K. P. Ferentinos, “Deep learning models for plant disease detection and diagnosis,” Computers and Electronics in Agriculture, vol. 145, pp. 311–318, 2018.
23. Geetharamani G., Arun Pandian J., Identification of plant leaf diseases using a nine-layer deep convolutional neural network, Computers & Electrical Engineering, Volume 76, Pages 323-338, ISSN 0045-7906, 2019.
24. Sharada P. Mohanty, David P. Hughes and Marcel Salathé, “Using Deep Learning for Image-Based Plant Disease Detection”, Sec. Technical Advances in Plant Science, Vol. 2, 2016.
25. Deng R, Tao M, Xing H, Yang X, Liu C, Liao K, Qi L. Automatic Diagnosis of Rice Diseases Using Deep Learning. Front Plant Sci. 2021 Aug 19; 12:701038. doi: 10.3389/fpls.2021.701038. PMID: 34490004; PMCID: PMC8416767.
26. Chen, J, Zhang, D, Suzauddola, M, Nanehkaran, YA, Sun, Y. Identification of plant disease images via a squeeze-and-excitation MobileNet model and twice transfer learning. IET Image Process. 2021;15: 1115– 1127. https://doi.org/10.1049/ipr2.12090
27. X. Zhang, Y. Qiao, F. Meng, C. Fan, and M. Zhang, “Identification of maize leaf diseases using improved deep convolutional neural networks,” IEEE Access, vol. 6, pp. 30 370–30 377, 2018.
28. A. K. Rangarajan, R. Purushothaman, and A. Ramesh, “Tomato crop disease classification using pre-trained deep learning algorithm,” Procedia Computer Science, vol. 133, pp. 1040–1047, 2018.
29. LeCun Y, Bengio Y, Hinton G., “Deep learning”, Nature, pp.436–44, 2015.
30. Sladojevic S., Arsenovic M., Anderla A., and Dubravko Culibrk D.S., “Deep neural networks based recognition of plant diseases by leaf image classification”, Computational Intelligence and Neuroscience, pp.1–12, 2016.

31. Pan, S.J.; Yang, Q. A survey on transfer learning. IEEE Trans. Knowl. Data Eng. 2010, 22, 1345–1359.
32. DeChant C., Wiesner-Hanks T., Chen S., and Stewart E. L., Automated Identification of Northern Leaf Blight-Infected Maize Plants from Field Imagery using Deep Learning, Phytopathology, Vol. 107(11), pp. 1426–1432, 2017.
33. Abu Sarwar Zamani, L. Anand, Kantilal Pitambar Rane, P. Prabhu, Ahmed Mateen Buttar, Harikumar Pallathadka, Abhishek Raghuvanshi, and Betty Nokobi Dugbakie, "Performance of Machine Learning and Image Processing in Plant Leaf Disease Detection", Journal of Food Quality, Hindawi, Article ID. 1598796, 2022.
34. R. Sangeetha, M. Mary Shanthi Rani, "A Novel Method for Plant Leaf Disease Classification Using Deep Learning Techniques", Machine Learning, Deep Learning and Computational Intelligence for Wireless Communication, 2021, Volume 7(4), 2021.
35. Deepalakshmi P., Prudhvi Krishna T., Siri Chandana S., Lavanya K. and Parvathaneni N, "Plant Leaf Disease Detection Using CNN Algorithm", International Journal of Information System Modeling and Design (IJISMD), Vol. 12(1), pp.1-21, 2021.