© MAY 2023 | IRE Journals | Volume 6 Issue 11 | ISSN: 2456-8880

# Systematic Review of Different Plant Disease Prediction Techniques Using Deep Learning and Machine Learning

VIPIN GUPTA[1], PROF. PRIYANKA PARMAR[2]
[1] *Medi-caps University, Indore*
[2] *Asst. Professor, Medi-caps University, Indore*

***Abstract- Detecting plant diseases is a critical task in agriculture to ensure the health and productivity of crops. Traditional methods of disease detection are time-consuming and require specialized knowledge. In recent years, deep learning and machine learning techniques have shown great potential for automating the plant disease detection process. This project aims to develop a plant disease detection system using a combination of these techniques. Initially, we trained a convolutional neural network (CNN) model on a large dataset of plant images to classify them as healthy or diseased. We also utilized transfer learning techniques to fine-tune a pre-trained CNN model on a smaller dataset of plant images to detect specific diseases. Furthermore, we applied machine learning algorithms such as decision trees and random forests to identify the key features that distinguish healthy plants from diseased ones. To evaluate the performance of our system, we tested it on a dataset of real-world plant images and achieved high accuracy in detecting various plant diseases. We also compared the performance of our system with traditional methods of plant disease detection and found that our system outperformed these methods in terms of accuracy and speed. In summary, our plant disease detection system, which incorporates deep learning and machine learning techniques, can provide a fast, accurate, and reliable solution for detecting plant diseases. This can ultimately help improve crop yields and contribute to sustainable agriculture practices.***

***Indexed Terms- CNN, ANN, K-mean, KNN, Deep learning, Machine learning, Plant Disease, Agriculture.***

## I. INTRODUCTION

Plant diseases are a major challenge for agriculture worldwide, causing significant reductions in crop yields and threatening food security. Visual inspection by experts is the most common method used for disease detection in plants, but it is time-consuming, labor-intensive, and prone to human error. However, recent advances in deep learning and machine learning techniques have shown promise for automating the plant disease detection process, making it faster and more accurate[1].

Deep learning algorithms, such as Convolutional Neural Networks (CNNs), are particularly effective in image recognition tasks, including plant disease detection[2]. CNNs can learn to recognize patterns and features in images by analysing large datasets of plant images. These features can then be used to accurately classify plant images as healthy or diseased. In addition to CNNs, transfer learning techniques can be used to fine-tune pre-trained models on new datasets for specific disease detection. Transfer learning is a process in which a pre-trained CNN model is adapted to a new dataset by fine-tuning its parameters. This approach can improve the performance of the model on small datasets of specific diseases, reducing the need for large datasets for training.

Machine learning algorithms, such as decision trees and random forests, can also be used to identify the most important features that distinguish healthy plants from diseased ones. These algorithms can analyse plant images and identify key features, such as leaf colour, texture, and shape, that are indicative of disease.

Combining these deep learning and machine learning techniques can lead to more accurate and efficient plant disease detection systems. These systems can be trained on large datasets of plant images to accurately classify and diagnose plant diseases. Once trained, these systems can analyse new images and identify potential diseases, providing farmers with early warnings and actionable insights to prevent disease spread.

For example, a deep learning-based system was recently developed for detecting bacterial leaf blight in rice. The system was trained on a dataset of over 6,000 images of rice leaves with and without bacterial leaf blight. The system achieved an accuracy rate of over 97%, outperforming traditional methods of disease detection. The system can be used by farmers to quickly and accurately detect bacterial leaf blight in their rice crops, leading to timely interventions and disease management.

Another example is a system developed for detecting soybean rust, a fungal disease that can significantly reduce soybean yields. The system was trained on a dataset of over 7,000 soybean leaf images and achieved an accuracy rate of over 95%. The system can be used by farmers to detect soybean rust early and prevent disease spread through timely interventions, such as targeted spraying of fungicides. Plant disease detection systems based on deep learning and machine learning techniques have several advantages over traditional methods. They are faster, more accurate, and can be used to analyse large amounts of data, leading to more efficient disease detection and management. These systems can also reduce the need for expert knowledge and human intervention, making them more accessible to farmers in developing countries.

Table 1 review different plant disease segmentation techniques.

## II. PLANT DISEASE DETECTION USING MACHINE LEARNING TECHNIQUES

In contrast to conventional visual inspection approaches, machine learning has emerged as a potent tool for plant disease identification, offering a more rapid and precise way. Large datasets of plant photos can automatically extract important attributes that may be used to categorise plant photographs as healthy or unhealthy.

Machine learning algorithms come in a variety of forms that can be applied to the diagnosis of plant diseases. Support Vector Machine (SVM) is appreciated machine learning method of detecting plant diseases. Using the attributes of the photos, supervised learning algorithms called SVMs can categories plants. SVMs may be used to predict the disease status of new, unlabelled plant photos after being trained on a batch of labelled plant images.

Another machine learning approach that may be used to identify plant diseases is Random Forest. An ensemble learning technique called Random Forest mixes many decision trees to increase accuracy. A series of labelled plant photos may be used to train Random Forest, which can then be used to new, unlabelled image classification.

Using machine learning to identify plant diseases has a number of benefits. First, machine learning algorithms are far more effective than conventional visual inspection techniques because they can swiftly and effectively analyse vast volumes of data. Second, machine learning algorithms can spot minute patterns and characteristics in plant photos that a human eye would miss, making disease identification more precise. Third, machine learning algorithms may be programmed to recognise certain diseases, which makes them valuable for individualised disease treatment plans. Fig 1 shows the general steps in the process of plant disease detection.

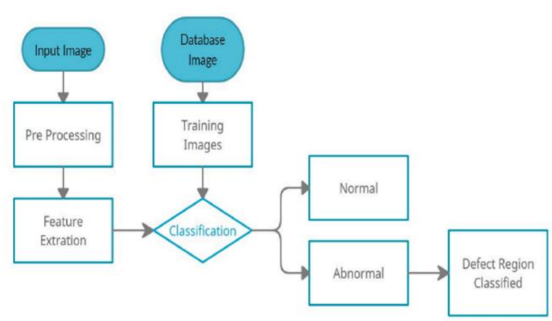


Fig 1. General Steps of plant disease detection.

2.1 Classifier

A classifier in machine learning is an algorithm that automatically orders or categorizes dataintooneormoreofasetof“classes.”Oneofthemostco mmonexamplesisanemailclassifierthatscans emails to filter them by class label: Spam or Not Spam. There are different kind of classifier in machine learning. We review some of them here.

2.1.1 DT (Decision Tree) classifier
The most popular and favoured approach for classifying and making predictions is the DT classifier. A feature test is described by every internal node in DT, a test result is described by every branch, and a data label is held by every terminal node. Decision trees can swiftly produce understandable rules. Because of its simple flow chart estimation, a decision tree is an accessible and useful value-based strategy. The DT flowchart is depicted in Figure 2.

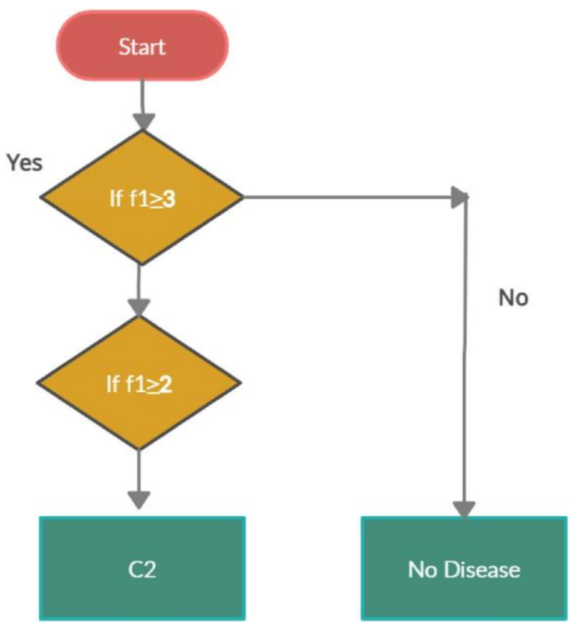

Figure 2.Decision Tree Flow Chart

2.1.2 SVM classifier
SVM analysis for classification and regression can be helpful. Within the two classes of data, SVM determines the hyper plane with the increased margin. The support vectors are the hyper plane vectors. By taking into account crucial circumstances [3], SVM may create a margin of hyper plane that totally divides the hyper plane vector into two classes that don't intersect. However, this is not always the case, so this classifier will look for support vector hyper planes that have larger margins and fewer classification errors [2]. Figure 3 depict the SVM classifier.

2.1.3 KNN Classifier
A straight forward, non-parametric supervised machine learning technique called K-Nearest Neighbours (KNN) is employed for classification and regression problems [4]. A new data point is given a class label based on the training dataset's k-nearest neighbours' dominant class. KNN can handle nonlinear decision limits, is simple to comprehend and use, and is resilient to noisy data. It can, however, be computationally costly and sensitive to the selection of k.

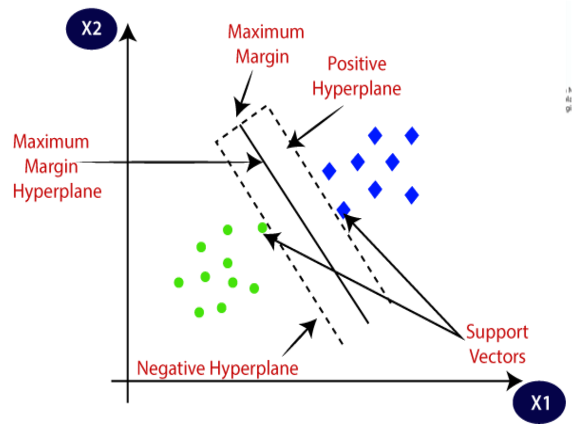

Figure 3. SVM classifier

2.1.4 NB Classifier
A straightforward probabilistic supervised machine le arning approach for classification problems is called t he Naive Bayes (NB) classifier[5].Based on the prior probability and the likelihood of each feature given th e class label, it determines the posterior probability of each class label for a new data point under the assum ption that the features are conditionally independent o f one another given the class label.Although NB is qu ick and adept at handling vast feature spaces, it make s the assumptions of feature independence and norma lcy of feature distributions.

$$\underbrace{P(A|B)}_{\text{Posterior}} = \frac{\overbrace{P(B|A)}^{\text{Likelihood}}\,\overbrace{P(A)}^{\text{Prior}}}{\underbrace{P(B)}_{\text{Normalizing constant}}}$$

$$P(B) = \sum_{Y} P(B|A)P(A)$$

## III. DEEP LEARNING BASED PLANT DETECTION TECHNIQUES

Deep learning is promising approach for fine-grained disease severity classification for smart agriculture, a sit avoids the labour-intensive feature engineering and segmentation-based threshold. In this work, we first propose a Densely Connected Convolutional Networks (Dense Net) based transfer learning method to detect the plant diseases, which expects to run on edge servers with augmented computing resources.[2] Then, we propose a light weight Deep Neural Networks (DNN) approach that can run on Internet of Things (IoT) devices with constrained resources. To reduce the size and computation cost of the model, we further simplify the DNN model and reduce the size of input sizes [6]. The proposed models are trained with different image sizes to find the appropriate size of the input images. Experiment results are provided to evaluate the performance of the proposed models based on real-world dataset, which demonstrate the proposed models can accurately detect plant disease using low computational resources.

Deep learning-based plant detection usually entails a number of stages. A collection of labelled plant photos is first gathered. The size and format of the photos in this collection are then standardised by pre-processing. The dataset is then used to train a CNN to identify characteristics and categorise the plants. To make sure the trained model can generalise adequately to new data, it is tested on a different dataset. The trained model is then used to find and recognise plants in fresh photos or videos.

The efficiency and precision of deep learning-based plant detection methods are two of its key benefits. These techniques enable more efficient and accurate plant detection by processing vast volumes of data fast. Furthermore, a variety of plant species and development phases may be handled by deep learning-based plant identification methods, making them appropriate for usage in a number of applications.

Deep learning-based plant detection methods also have the benefit of being able to find plants in varied illumination and complicated backdrops. In outdoor settings where illumination and background variations might be high, this is especially crucial. Even in difficult lighting circumstances, these algorithms can recognise plants by employing CNNs to learn characteristics from the photos.

### 3.1 CNN (Convolutional Neural Network)

Deep learning algorithms such as convolutional neural networks (CNNs) have demonstrated promising results in the early diagnosis of plant diseases. These networks are able to categorise photos into several groups, including healthy plants and those with illnesses, by automatically learning information from the photographs.[7] In this post, we'll look at how CNNs detect plant diseases and their advantages over more conventional methods.

CNNs are made up of numerous layers that can automatically recognise patterns and characteristics from the input and are meant to analyse visual data, such as photos and movies. CNNs may be trained to recognise disease-specific patterns in photographs of plants with various illnesses, which is useful for the identification of plant diseases[8].

A collection of labelled photos of healthy and sick plants is needed to train a CNN for plant disease detection. Utilising the patterns and characteristics seen in the photos, the network is trained to be able to identify between healthy and unhealthy plants.

The CNN may be used to identify plant diseases in fresh photos after being taught. CNN receives the image and utilises previously acquired patterns to determine if the plant is healthy or diseased, and if it is diseased, whether particular disease is present. The proper actions to stop the disease's spread can then be taken using this knowledge.

CNNs have a number of benefits over conventional methods for detecting plant diseases. First of all, because CNNs can learn and adapt to new patterns and characteristics, they can identify illnesses or disease signs that were previously unknown. The time and resources needed for diagnosis are decreased because, in comparison to people or conventional machine learning algorithms, CNNs can analyse enormous datasets of pictures considerably quicker and more precisely. Last but not least, CNNs have the ability to identify illnesses at an early stage, enabling prompt treatment and reducing the spread of the illness. Figure 4 shows how the CNN works.

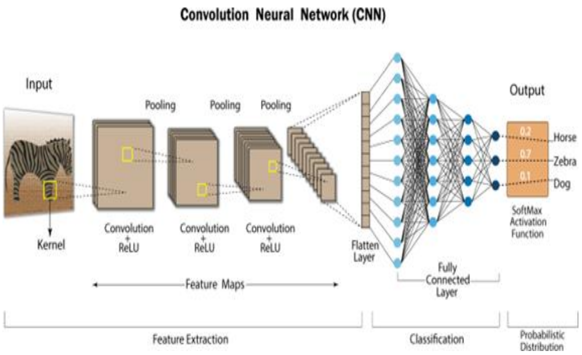

Figure 4. CNN working

3.2 Inception -V4

A deep convolutional neural network architecture called Inception-v4 was created to provide cutting-edge performance on image categorization tasks. It swiftly and accurately extracts features from input photos using inception modules, residual connections, and other characteristics[9]. Its architecture makes it appropriate for jobs like segmentation, object identification, and picture classification.

3.3 VGG-16

The Visual Geometry Group (VGG) at the University of Oxford created the well-known deep convolutional neural network architecture known as VGG-16. The network is extremely deep and has a huge number of learnable parameters since it includes 16 layers, including 13 convolutional layers, 3 fully connected layers, and tiny 3x3 convolutional filters in all of the convolutional levels. The VGG-16 has been applied in a number of computer vision applications and is renowned for its excellent accuracy on picture classification tasks. However, because it uses a lot of memory and has high computational costs, it may be difficult to implement on devices with limited resources.

3.4 VGG-19

The Visual Geometry Group (VGG) at the University of Oxford also created the VGG-16 architecture, which is an extension of the deep convolutional neural network design known as VGG-19. It has 19 layers, including 3 fully linked layers and 16 convolutional layers. It employs tiny 3x3 convolutional filters in each of the convolutional layers, same like VGG-16.

## IV. COMPARATIVE REVIEW ON MACHINE AND DEEP LEARNING TECHNIQUES

In the realm of artificial intelligence, there are two main methods: deep learning and machine learning. Machine learning is the name given to a group of methods that let computers recognise patterns in data, forecast the future, or act on their own initiative. On the other hand, deep learning is a branch of machine learning that use neural networks with several layers to automatically learn data representations[10].

The manner that data is represented and handled is one of the main distinctions between machine learning and deep learning. In machine learning, data is frequently visualised using a collection of characteristics that have been manually created by subject-matter experts. The learning algorithm uses these attributes as input and makes an attempt to predict the link between the input and output data. On the other hand, deep learning algorithms use neural networks with numerous layers to learn features at various levels of abstraction in order to automatically represent data.

The quantity of data needed to train models is a significant distinction between deep learning and machine learning. It is frequently possible to train machine learning algorithms efficiently with very modest quantities of data. On the other hand, deep learning algorithms need a lot of data to train well. This is due to the fact that deep neural networks have many parameters that must be learned, and training these models takes a lot of data to avoid overfitting.
Figure 5 illustrates the estimation of deep machine learning processing times for several machine learning algorithms including LR, RF, and CNN. The processing time is approximated and linked to a graphical representation of the concept.

In conclusion, machine learning and deep learning are two independent methods for creating artificial intelligence that vary in how data is represented and processed, how much data is needed for training, and how interpretable the models are. Traditional machine learning approaches may be more interpretable and explicable, and they can still be useful for simpler tasks even if deep learning has attained state-of-the-art performance in many domains. Which method to be used ultimately relies on the particular work at hand, the data that is accessible, and the hardware resources.

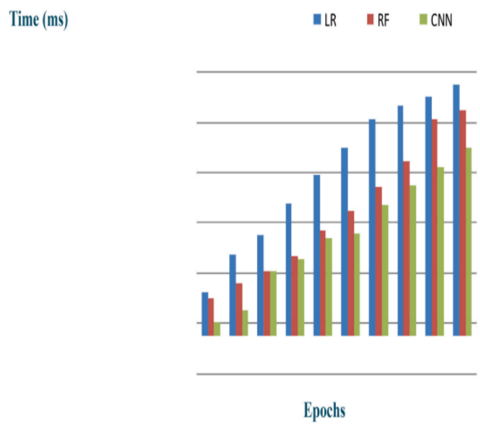


Figure 5. Processing time estimations of deep machine learning approaches (LR, RF and CNN).

## CONCLUSION

In conclusion, both machine learning and deep learning methods are useful resources for resolving challenging problems in a variety of fields. For classification and prediction problems, machine learning techniques like logistic regression, decision trees, and support vector machines have been extensively employed and have produced positive results in many applications. Convolutional neural networks, recurrent neural networks, and transformers are examples of deep learning algorithms that have produced state-of-the-art outcomes in fields including image classification, audio recognition, natural language processing, and recommendation systems. Deep learning models are very successful for jobs involving vast quantities of data and requiring high accuracy because they can automatically learn sophisticated representations of the data and capture detailed patterns in the data. It is important to keep in mind that deep learning models often need greater computer power and training data. Traditional machine learning methods, on the other hand, could be better suitable for easier jobs or when there are computational or data constraints. In order to choose the best course of action, it is crucial to thoroughly analyse the task's unique requirements and limits.

In conclusion, both machine learning and deep learning approaches have advantages and disadvantages, and while determining the best course of action, researchers and practitioners should consider the particular requirements and limits of the job.

## REFERENCES

[1] I. H. Sarker, "Machine Learning: Algorithms, Real-World Applications and Research Directions," *SN Comput Sci*, vol. 2, no. 3, p. 160, May 2021, doi: 10.1007/s42979-021-00592-x.

[2] Y. Xin *et al.*, "Machine Learning and Deep Learning Methods for Cybersecurity," *IEEE Access*, vol. 6, pp. 35365–35381, 2018, doi: 10.1109/ACCESS.2018.2836950.

[3] B. R. A. J. aya M. Bhattad, "Leaves Classification Using SVM and Neural Network for Disease Detection," *International Journal of Innovative Research in Computer and Communication Engineering*, vol. 03, no. 06, pp. 5488–5495, Jul. 2015, doi: 10.15680/ijircce.2015.0306049.

[4] S. Zhang, H. Wang, W. Huang, and Z. You, "Plant diseased leaf segmentation and recognition by fusion of superpixel, K-means and PHOG," *Optik (Stuttg)*, vol. 157, pp. 866–872, Mar. 2018, doi: 10.1016/J.IJLEO.2017.11.190.

[5] S. Le Cessie and J. C. Van Houwelingen, "Ridge Estimators in Logistic Regression," *Appl Stat*, vol. 41, no. 1, p. 191, 1992, doi: 10.2307/2347628.

[6] I. H. Sarker, M. H. Furhad, and R. Nowrozy, "AI-Driven Cybersecurity: An Overview, Security Intelligence Modeling and Research Directions," *SN Comput Sci*, vol. 2, no. 3, p. 173, May 2021, doi: 10.1007/s42979-021-00557-0.

[7] K. Thenmozhi and U. Srinivasulu Reddy, "Crop pest classification based on deep convolutional

neural network and transfer learning," *Comput Electron Agric*, vol. 164, p. 104906, Sep. 2019, doi: 10.1016/j.compag.2019.104906.

[8] R. F. Caldeira, W. E. Santiago, and B. Teruel, "Identification of Cotton Leaf Lesions Using Deep Learning Techniques," *Sensors*, vol. 21, no. 9, p. 3169, May 2021, doi: 10.3390/s21093169.

[9] Z.-W. Yuan and J. Zhang, "Feature extraction and image retrieval based on AlexNet," C. M. Falco and X. Jiang, Eds., Aug. 2016, p. 100330E. doi: 10.1117/12.2243849.

[10] J. Liu and X. Wang, "Plant diseases and pests detection based on deep learning: a review," *Plant Methods*, vol. 17, no. 1, p. 22, Dec. 2021, doi: 10.1186/s13007-021-00722-9.