International Journal of Science and Engineering Applications
Volume 12-Issue 05, 01 - 06, 2023, ISSN:- 2319 - 7560
DOI:10.7753/IJSEA1205.1001

# A Comprehensive Survey on Plant Leaf Disease Detection Using Image Analytics

Sanika Singh
Banasthali Vidyapith
Rajasthan, India

Saurabh Mukherjee
Banasthali Vidyapith
Rajasthan, India

**Abstract:** A country's economy depends heavily on agriculture, and there are many varieties of crops that can be cultivated by farmers. The issue or problems arise when the farmers are unaware of a plant disease that is affecting their crops at the appropriate moment, thus harvests become afflicted. Most of the time farmers are not aware about the disease and its type when it is discovered.Given these challenges, research in the field of automatic leaf disease detection in agriculture is of significant importance, since it could offer benefits in the detection of bigger fields of crops and help to detect diseases since they appear on plants leaf. The examination to various pattern present above the plant leaves is necessary for study of plant disease.

**Keywords:** Image Analytics, Machine Learning (ML), Artificial Neural Network (ANN), Image Segmentation, Automatic disease Detection.

## I. INTRODUCTION

In agriculture, the soil used to be plow over, which encourages the growth of plants and a variety of vegetables and fruits. In order for people to live happily on earth and continue to obtain food and other necessities like food, milk, wool, and a variety of other things, plants must be present in our environment and support animal raising.

Agriculture is dependent on the number and quality of agricultural products, especially plants, in every nation. Plant diseases cause major financial losses in agricultural output.

It's crucial to quickly identify and classify plant illnesses in order to treat and monitor them. Effective plant disease defense has a strong relationship with sensible horticulture and environmental changes.

### *A. Convolutional neural networks (CNN)*

For the purpose of image recognition as well as computer vision activities,a type of Artificial Neural Network called Convolutional Neural Network (CNN) is extremely important. There are various layers in the CNN architecture. Layering for convolution Filtering is done at the convolution layer.

**Pooling Layer:** Down sampling is another name for it. As implied by the name, it lessens the amount of data from the convolution layer is used in each feature. It simply keeps information that is necessary for the user. There are several rounds of the pooling and convolution layers so that we only have exact data [16]. The function of this layer is shown in fig. 1.

Flattens is another name for the fully connected input layer. In a neural network, a fully connected layer takes input from the previous layer and combines it into a single vector [10]. This vector serves as the output of the layer and is passed on as input to the next layer in the network. The final fully connected layer produces a probability score for each potential output class, while the first fully connected layer uses the applied weights to predict the correct label.

### *B. Artificial Neural Network (ANN)*

The artificial neural network used in machine learning is inspired by the human nervous system, or the brain. There are many hidden layers in this, the output layer provides the output. The input is processed at input layers. Between the input-output layers there are hidden layers that apply weights to the inputs via functions and route the output through an activation function [2]. ANNS are used for a variety of machine learning tasks, such as classification, regression, and prediction. They have been successful in many applications, including image recognition, natural language processing, and speech recognition.

During training, the network is given examples of input-output pairs and the weights are adjusted to minimize the difference between the predicted output and the actual output as shown in fig. 2. This is typically done using an optimization algorithm, such as gradient descent, which adjusts the weights in a way that minimizes the error between the predicted output and the actual output. Once the network has been trained, it can be used to make predictions on new input data. The input data is processed through the network, and the output is produced by the final layer of neurons [3]. The accuracy of the network's predictions can be evaluated by comparing the predicted output to the actual output for a set of test data.

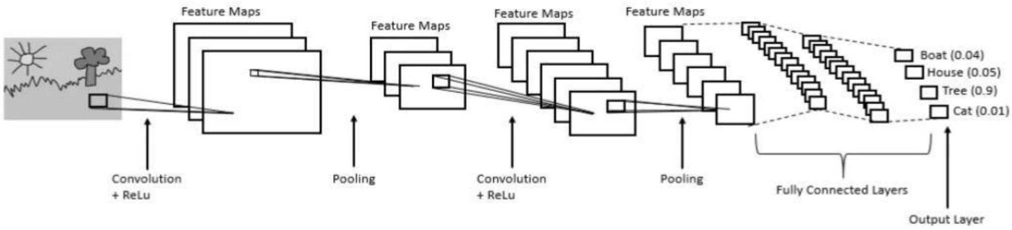


Fig.1: CNN architecture.

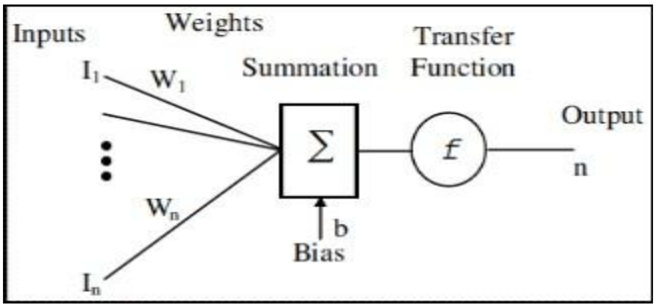


Fig. 2: ANN working principle.

*C. Support Vector Machine (SVM)*

SVM is one of the usual supervised machine learning approaches which is used the most for classification and regression issues. But this classifier is mainly used in machine learning to solve categorization issues. The SVM algorithm is used for binary classification problems where there are two groups. It aims to create an optimal decision boundary or line in ndimensional space that can accurately separate the groups, making it easier to classify new data points in the future. This decision boundary is referred to as a hyperplane [13]. The SVM algorithm identifies extreme data points, known as support vectors, that are used to build the hyperplane. The process is therefore referred to as support vector machines. This is illustrated in figure 3.

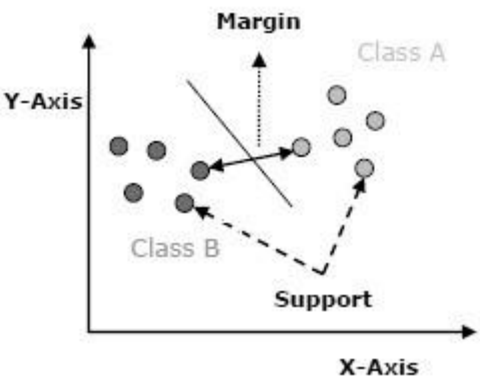


Fig. 3: SVM working principle.

*D.Generative Adversarial Networks (GANs)*

This group of frameworks for machine learning. It has been discovered that Lan Goodfellow et.al, Generative Adversarial Networks, which they developed in 2014, are effective in a range of problems requiring the development of synthetic images. The primary goal is to generate artificial data that matches to the training distribution's criteria as closely as possible. The conversion of one visual representation of a scene to another, known as image transformation, utilizes GANs that are driven by the achievement of a specific outcome.

In simpler terms, a Generative Adversarial Network (GAN) is a system that consists of two neural networks, a generator and a discriminator, working against each other to produce synthetic data. The generator network accepts random noise as input and generates artificial/synthetic data, while the discriminator network receives both real and synthetic data and attempts to differentiate between them. [1].

The two networks are trained simultaneously, with the generator trying to produce data that can fool the discriminator, and the discriminator trying to correctly identify the synthetic data. This process continues until the generator is able to produce data that is indistinguishable from real data, resulting in the production of synthetic data that can be used for various tasks.

Generative Network learns about producing features during the training phase, which it subsequently imparts to the Discriminative Network for the relevant category. This is how GAN works.

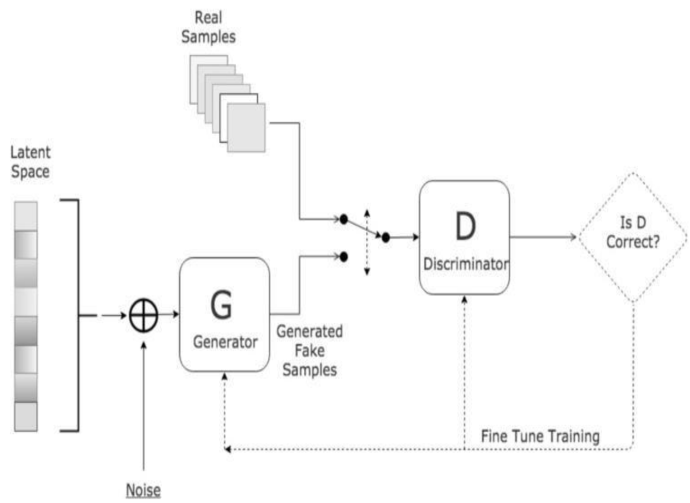


Fig. 4: Working of GAN.

International Journal of Science and Engineering Applications
Volume 12-Issue 05, 01 - 06, 2023, ISSN:- 2319 - 7560
DOI:10.7753/IJSEA1205.1001

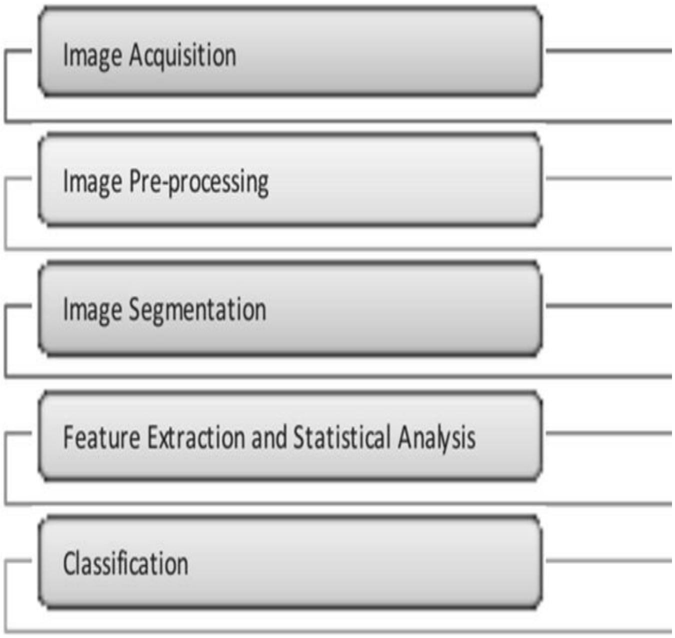


Fig.5: Disease detection steps.

a. **Image acquisition:** To take pictures of the leaves at this stage, we utilized a digital camera or a phone.
b. **Image Preprocessing:** After gathering images, we use image processing methods to clean up any poor quality or unclear regions of interest in the image.
   Some image preprocessing methods include cleaning, integrating, transforming, and reducing.
c. **Image segmentation:** It is a process of breaking an image up into pieces that may be compared or have characteristics in common. By using image segmentation, we are able to isolate the specific area of the image that we need while maintaining its original properties.
d. **Feature Extraction & Statistical Analysis:** For the purpose of identifying an object, feature extraction is essential. The picture processing feature is used in a variety of situations. Plant diseases can be identified using various characteristics, such as color, texture, brightness, shape, and edges. To make it easier for algorithms to process, these traits are condensed into a simpler set of features. This process of simplifying the input data is called feature extraction. This is done because the input data is often too large and contains too much repetition for algorithms to handle.
e. **Classification:** Utilizing training data, the Classification method is a supervised learning method for classifying new observations. After learning from the provided dataset or expectations of categorization, a computer program classifies new observations into multiple classes or categories.

Groups can also be referred to as goals, labels, or divisions. Two categories of algorithms for classification exist: Support Vector Machines and Logistic Regression are two subcategories of **linear models**. KNN, Naive Bayes, Random Forest, and SVM Kernel, these are examples of **non-linear models**.

## II. LITERATURE SURVEY

**Nazki et al. [1]** employed Deep Convolution Neural Network after processing images using the GAN technique to improve the performance of plant disease detection. Due to the sample image generated by GAN's restricted and practical features, it was able to advance further.

According to the authors, incorporating synthetic samples generated by GAN architecture resulted in a performance improvement of 5.2%, which was higher than the 0.8% improvement observed with traditional augmentation techniques. Another study by **Ganatra and Patel[2]** proposed a convolutional neural network approach for plant disease classification using images. The study compared various frameworks including VGG16, Inception V4, ResNet 50, and ResNet 101. The evaluation showed that ResNet50 and ResNet101 had accuracy rates of 99.70% and 99.73%, respectively.

Similarly, **Thushara et.al. [3]** offered a unique form of deep learning system that uses images of leaves to swiftly identify and analyze plant illnesses. The authors of the study recommend a technique that is able to distinguish between four different types of diseased leaves and healthy leaves. They were able to achieve an accuracy of 96% by using real-time photographs of both diseased and healthy leaves.

With the development of creativity, tracking control and the executive framework are being used more and more, according to **Chaitanya and Yasudha [4].** The vast reach of disease is mostly to blame for horticulture crop death. The model has a 98.84 percent accuracy rate. A novel deep learning framework called GPDCNN was given by **Zhang et al. [5].** For the detection of cucumber leaf disease. This increases the convolution receptive field by using global pooling layers in place of completely linked layers. Six set of cucumber disease image make up the dataset, and the result achieved is 94.65% accurate.

Additionally, Multilayer Convolutional Neural Network model was developed for detecting Anthracnose, a fungal disease in mango leaves, using real-time images by **Singh Chouhan et. al. [6]**. Images of affected and non-affected leaves are included in the dataset.

Using this model, 97.13% accuracy was attained. Additionally, the image segmentation method that **Vijai Singh [7]** suggested for identifying disease in sunflower plant leaves was successful in detecting and coordinating the infections. The proposed algorithm stands out as it doesn't rely on prior knowledge or require a specific amount of data to train. It is more efficient compared to other existing methods. The algorithm has a typical classification accuracy of 98.0%.

A Deep learning model was also developed **Mishra et.al, [8]** for the purpose of detecting illness in corn leaves. The detection rate for illness in corn leaves was 88.66%.

Additionally, **Sharma et al. [9]** the CNN model looked at a typical outcome to the issue utilizing fragmented image data. Execution of the model went up from 42.3% to 98.6%. In moreover, an increase in conviction was seen in 82% of the test datasets for the quantitative study of self-agreement conviction.

**Agarwal et.al, [10]** developed a convolutional neural network (CNN) based model for the detection of tomato leaf diseases. The model is composed of three convolution layers and the maximum number of pooling layers. The dataset, which includes nine kinds of both sick and healthy plants, is from Plant Village.

**Khamparia et.al, [11]** a deep convolution encoder network model was created. This design used in plant leaf disease identification. The dataset used in the research is sourced from Plant Village, it includes both healthy and diseased leaves of corn plant. The algorithm developed using this dataset was able to correctly identify 97.50% of cases of corn leaf diseases.

**Venkataramanan et al. [12]** demonstrated a method using deep learning to identify and classify plant diseases based on the leaves of certain plants. To improve forecast accuracy, the categorization process was carried out in stages, excluding probable outcomes at each stage.

**Sampoorna and Rasadurai [13]** used a combination of different techniques, such as K-Means clustering, Otsu Segmentation, a Convolutional Neural Network, and a Support Vector Machine classifier algorithm to identify and classify various characteristics of a plant. They also utilized various sensors like DHT11, Soil PH, Soil Moisture, and UV sensor to monitor and control the growth of the plant by interfacing it with a microcontroller.

This method produced dependable results and was able to produce results with more precision than the multilayer Perceptron procedure. The architectural precision of the system is 11.0, that is 0165% greater from the current system accuracy. **Hussain et al. [14]** states that deep learning approach was created for the identification of wheat diseases based on on-site images captured by cameras with different settings. The dataset consisted of four types of wheat diseases, namely stem rust, yellow rust, powdery, and normal, with each category containing a total of 2,207 images.

**Sladojevic et al. [15]** employed a deep convolution network to the most recent generation of convolution neural networks to address the unique method for plant disease diagnostics and saw amazing results in the classification of images. The model was capable of differentiating between 13 various types of specific disorders. DCNN was trained using a deep learning framework developed by the Berkeley Vision and Learning Center. Which helped in achieving an average precision of 96.3% with a range from 91% to 98%.

To track various pests in coconut plants as it passes over the coconut orchard, **Abraham Chandy [16]** also used a camera-interfaced robot and the NVIDIA Tegra Machine on Chip (SoC), a precision farming technique. The authors used a drone to gather images of trees, which were then analyzed using a deep learning algorithm to identify trees affected by disease or pests. The photos were processed using a machine learning algorithm to detect the diseased and pest-affected trees. The deep learning algorithm was used to improve the accuracy of identification. The farmer's mobile phone received the info right away thanks to Wi-Fi. This makes it easier to control pest-infested trees quickly and increases tree yield.

**Vinothkanna and Vijayakumar [17]** developed a RESNET 152 deep Convolutional Neural Network (CNN) based model for identifying the various stages of Dragon fruit. They used live images for testing the model. Their results showed that the model had a higher accuracy in comparison to the VGGNET model even when the number of epochs were increased. The Area Under the Receiver Operating Characteristic Curve (AUROC) was 1, which indicates the model has a high performance in identifying the stages of the dragon fruit.

**Table 1: ReviewSummary**.

| Paper | Techniques used | Pros | Cons |
|---|---|---|---|
| Multi-layered Convolution neural network for the Classification of mango leaves infected by Anthracnose Disease. | Multilayer convolutional neural network (MCNN) | It can identify important features without any human input. | Has multiple layers, and its training process can take a significant amount of time if the computer has a weak CPU. |

| | | | |
|---|---|---|---|
| A detection system based on a Deep Convolutional Neural Network was developed to recognize and identify corn plant diseases in real-time. | Deep Convolution Neural Network | Less dependent on preprocessing, which reduces the need for human effort. It is capable of learning on its own, which simplifies the pre-processing phase | It is a self-learning algorithm, which simplifies the preprocessing stage. |
| Convolutional Neural Network (CNN) utilized for detecting tomato leaf diseases. | Convolution Neural Network | The proposed model requires a storage space of approximately 1.5MB, while pre-prepared models require an additional 100MB. | A CNN may be slower due to certain operations such as pooling. |
| Recognition of plant diseases through leaf image classification was achieved using | Deep Convolution Neural Network | DCNNs typically have many layers, with the number of layers ranging from several to dozens. The use of more layers can | CNNs are not able to take into account the position and orientation of an object in an image, as they |
| Deep Neural Networks. | | lead to increased accuracy, as it allows the network to learn more complex features of the input data. | designed to be spatially invariant. |
| The identification of cucumber leaf diseases using a Global Pooling Dilated Convolutional Neural Network. | GPDCNN | The ability to perform better than other strategies in terms of robustness. | |
| Detection of sunflower leaf diseases through Image Segmentation using Particle Swarm Optimization. | Particle Swarm Optimization Algorithm. | One of the key advantages of PSO is that it requires minimal user intervention and has few parameters that need to be adjusted, which makes it easy to use. | The algorithm's basic nature may not be sufficient to effectively solve more complicated issues, making it less suitable for certain types of problems. |

## III. CONCLUSION

In this article, we discussed the fundamental methods utilized by different researchers to identify plant diseases.The tabular analysis method offers techniques for identifying, dividing and categorizing data based on different datasets and their specific characteristics. The GPDCNNs have a higher recognition and learning capability than other methods. Visual representations of information appropriation in GANs are more efficient (images that are more refined and clearer). Multilayered Convolutional Neural Network has the benefit of being able to recognize significant aspects without human supervision because it departs from its paradigms. In other words, we have presented a summary of the different methods used to address the problem at hand, as well as the datasets that support these methods.

## IV. FUTURE WORK

Integrating computer vision with techniques for identifying plant diseases is our long-term goal. We wish to increase GPDCNN and Convolution Neural Network capabilities going forward and detect different diseases. Also work can be done to do the real time detection that can be useful in controlling the spread of plant diseases. More number of classes can be involved in the study which can detect maximum possible diseases in plants.

## REFERENCES

1. Haseeb Nazki, Sook Yoon et.al., (2020) “Unsupervised image translation using adversarial networks for improved plant disease recognition” Published by Elsevier B.V.
2. Nilay Ganatra, Atul Patel (2020) “Performance Analysis of Fine-Tuned Convolutional Neural Network Models for Plant Disease Classification” Published by International Journal of Control and Automation Vol. 13, No. 3, pp. 293-305.
3. Battala lakshmi thushara et.al., (2020) “Analysis of Plant Diseases Using Expectation Maximization Detection with BP-ANN Classification” Volume XIII, Issue VIII, AUGUST 2020.
4. Chaitanya and Yasudha (2020) “Image Based Plant Disease Detection Using Convolution Neural Networks Algorithm”, Volume 5, Issue 5, May – 2020, Published by International Journal of Innovative Science and Research Technology ISSN No:-2456-2165.
5. Shanwen Zhang, Subing Zhang et.al., (2019) “Cucumber leaf disease identification with global pooling dilated convolutional neural network” 2019 Published by Elsevier B.V.
6. Uday pratap singh, et.al. (2019) “Multilayer Convolution Neural Network for the Classification of Mango Leaves Infected by Anthracnose Disease” 2019.
7. Vijai Singh (2019) “Sunflower leaf diseases detection using image segmentation based on particle swarm optimization” 2019 Published by Elsevier.
8. Sumita Mishra, Rishabh Sachan, Diksha Rajpal (2019) “Deep Convolutional Neural Network based Detection System for Real-time Corn Plant Disease Recognition” 2020 Published by Elsevier B.V.
9. Parul Sharma, Yash Paul Singh Berwal , Wiqas Ghai (2019) “Performance analysis of deep learning CNN models for disease detection in plants using image segmentation” open access 2019 Published by Elsevier B.V.
10. Mohit Agarwal, Abhishek Singh, Siddhartha Arjaria, Amit Sinha, Suneet Gupta (2019) “Leaf Disease Detection using Convolution Neural Network” 2019-2020 Published by Elsevier.
11. Aditya Khamparia, Gurinder Saini, Deepak Gupta, Ashish Khanna, Shrasti Tiwari, Victor Hugo C. de Albuquerque (2019) “Seasonal Crops Disease Prediction and Classification Using Deep Convolutional Encoder Network” 2019.
12. Aravindhan Venkataramanan, Deepak Kumar P Honakeri, Pooja Agarwal (2019) “Plant Disease Detection and Classification using Deep Neural Networks”, p-ISSN: 2229- 5631 Vol. 11 No. 08 Aug 2019.
13. C.K Sampoorna and Dr. K Rasadurai (2019) “Tomato Plant Leaf Disease Detection Using K-Means and SVM Classifier”, June 2019, Volume 6, Issue 6 www.jetir.org (ISSN-2349-5162), Published by Journal of Emerging Technologies and Innovative Research (JETIR).
14. Altaf Hussain, Mohsin Ahmad, Imran Ahmad Mugha, Haider Ali (2018) “Automatic Disease Detection in Wheat Crop using Convolution Neural Network” The 4th International Conference on Next Generation Computing 2018.
15. Srdjan Sladojevic, Marko Arsenovic, Andras Anderla, Dubravko Culibrk and Darko Stefanovic (2016) “Deep Neural Networks Based Recognition of Plant Diseases by Leaf Image Classification”, Volume 2016 Hindawi Publishing Corporation.
16. Dr. Abraham Chandy “Pest Infestation Identification in Coconut Trees Using Deep Learning” Journal of Artificial Intelligence 1, no. 01 (2019): 10-18.
17. Vijayakumar, and Mr R.Vinothkanna “Mellowness detection of dragon fruit using deep learning strategy” Journal of Innovative Image Processing (JIIP)2. no.01 (2020): 35-43.