

A SURVEY PAPER ON PLANT DISEASE IDENTIFICATION USING MACHINE LEARNING

¹ Dr Shankar Gowda B B , ² Karthik v Karanam

¹ HOD, Professor, Department of Master of Computer Application, BIET, Davangere

² Student, Department of MCA, BIET, Davangere

Abstract: Agriculture plays an important role in farmer's life. Sometimes manual identification of disease is time consuming and need of labor is more. One of the most important facts that reduce the growth of plants is disease attack. Overall study about agriculture shows that quality and quantity of agricultural products may be reduced due to various factors of plant diseases. These diseases can be more easily identified by using machine learning approach as compared to manual method. Hence machine learning method can be used to identify the affected leaf images. Images captured by camera will be processed using different image processing technique. These techniques will help in identifying plant diseases thereby increasing the yield of plants. This survey paper describes plant disease identification using Machine Learning Approach and study in detail about various techniques for disease identification and classification is also done.

Keywords: Feature extraction, image processing techniques, SVM, PNN, ANN and GA

I. Introduction

Agriculture is one of the important source of income for farmer. Farmers can grow variety of plants but diseases hamper the growth of plants. One of the major factors that leads the destruction of plant is disease attack. Disease attack may reduce the productivity plants from 10%-95%. [1] At present there are different strategies to get rid of plant diseases such as removing the affected plants manually, mechanical cultivation and last is using different pesticides. The easy method to detect to plant disease is taking help of agricultural expert. But this method of manual detection of diseases takes lot of time and is laborious work. Next method is using pesticide but excess use of pesticide may increase growth of plants but it reduces the quality of plant. But using more pesticide for plants without analysing how much quantity of pesticide is needed for particular crop because excess use of pesticide may lead adverse effect on environment and human health. Classification of Plant and Diseased Plants using digital image processing and Machine Learning approach which can help to control growth of

diseases on Plants using the pesticides in the quantity needed so that excess use of pesticides can be avoided.

Automatic identification of plant diseases is an important task as it may be proved beneficial for farmer to monitor large field of plants, and identify the disease using machine learning approach. As compared to image processing technique using machine learning approach manual disease identification is less accurate and time consuming.



Fig1.1: Diseased Plant leaves

In figure 1.1 we can notice different images of diseased plants this disease can be easily identified using machine learning and image processing techniques instead of manually.

II. Literature Survey

In order to know in detail about this survey the previous research work done in this direction, several studies dedicated to the topic were referred. The literature survey is done in chronological order from 2012-2016.

F Ahmed in 2012, [1] has stated that in most agricultural systems, one of the major concerns is to reduce the growth of Unwanted Plants. In most cases, removal of the Unwanted Plants population in agricultural fields involves the application of chemical herbicides, which may lead to huge productivity but adverse effect on environment and human health. The

ability of locating and classifying plants and Unwanted Plants in digital images could lead to development of autonomous vision guided agricultural equipment's for site-specific herbicide application.

Suhaili Beeran Kutty in 2013, [2] have considered an artificial neural network (ANN) based system to classify the watermelon leaf diseases of Downey Mildew and Anthracnose. This classification is based on the colour feature extraction from RGB colour model which is obtained from the identified pixels in the region of interest. Result of this work showed that the leaf disease achieved 75.9% of accuracy based on its RGB colour component

Godliver Owomugisha in 2014, [3] has stated that Machine learning is one of the approach which has been applied in Agriculture but also in other fields including crop disease detection for some crops.

Mrunmayee Dhakte in 2015, [4] has stated that Pomegranate is the one of the fruit grown in large quantity in many states of India and one of the most profit gaining fruit. But due to various reasons plants are infected by various diseases which destroys the crop. The work proposes an image processing and neural network methods to deal with main issues of phytopathology. In this system GLCM feature extraction techniques and k-means machine learning approach are used. Accuracy of this system to detect disease is 90%.

Arti Singh in 2016, [5] has stated that Advances in automated and high throughput images has resulted in deluge of high-resolution images and sensor data of plants. Extracting patterns and features from this set of data requires the use of machine learning techniques to find out data assimilation and feature identification. Different steps in which machine learning approaches can be deployed are Identification, classification, quantification and prediction.

Sharada P. Mohanty in 2016, [6] has stated that crop diseases are the main threat to food security, but their rapid detection remains difficult in many parts of the world due to lack of required infrastructure. The combination of increasing global use of smartphone and recent advances in computer advances made possible by deep learning has paved the way for smartphone assisted disease diagnosis. The trained model achieves 99.35% accuracy.

III. Machine Learning Approaches

Machine learning is field of computer science that uses statistical techniques to give computer system the ability to "learn" with data, without being explicitly programmed.

3.1 Decision Tree Learning: Decision tree comprises a set of rules that provide the means to associate specific molecular features and/or descriptor values with the region of interest.

3.2 Association Rule Learning: Association rule learning is a method for discovering interesting relations between variables in large database.

3.3 Artificial Neural Network: Artificial Neural Network (ANN) also called as neural networks inspired by biological neural networks.

3.4 Support Vector Machines: Support Vector Machines (SVMs) are set of related supervising methods used for classification and regression.

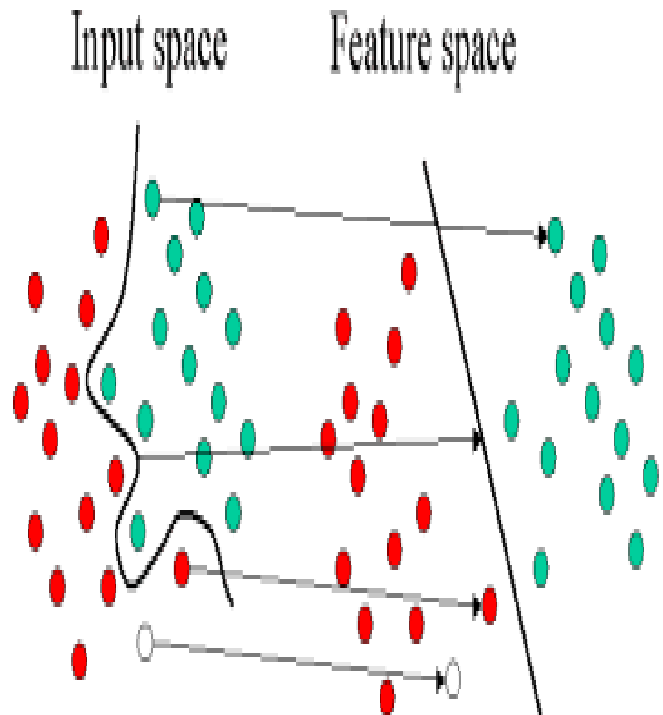


Fig: SVM algorithm

3.5 k-Nearest Neighbour: k-Nearest neighbour is pattern recognition algorithm. k-Nearest Neighbour is a simple classifier in the machine learning techniques where the classification is achieved by identifying the nearest neighbours to query examples and then make use of those neighbours for determination of the class of query. k-Nearest neighbour is type of lazy learning. A commonly used distance metric for k-NN is Euclidean distance.

IV. Comparison of Machine Learning Techniques

Table: Comparison between machine learning techniques

Machine Learning technique	Advantages	Disadvantages
Decision Tree Learning	<ul style="list-style-type: none"> Ability of selecting the most discriminatory features Handling both continuous and discrete data 	<ul style="list-style-type: none"> They are unstable Calculations can get very complex as time goes on
Association Rule Learning	<ul style="list-style-type: none"> Finds similar patterns from data and produces rules. Generates association relationship from data. 	<ul style="list-style-type: none"> Helps to find sequential patterns Uses acquisition, integration and integrity checks methods..
Support Vector Machine	<ul style="list-style-type: none"> Supervised Learning Technique Helps for classification and prediction purpose 	<ul style="list-style-type: none"> Produces very accurate classifiers Less over fitting and handles noise
Clustering	<ul style="list-style-type: none"> The objectives of the technique are as follows: <ul style="list-style-type: none"> ✓ To discover natural groupings ✓ To produce hypothesis from data ✓ To find reliable organization of data. 	<ul style="list-style-type: none"> Provides end user high level view of what is going in database. Very efficient technique
k-Nearest Neighbor	<ul style="list-style-type: none"> Robust to noisy training data Effective if training data is large 	<ul style="list-style-type: none"> Need to determine value of parameter of k Computation cost is high
Genetic Algorithm	<ul style="list-style-type: none"> It can find fit solutions in less time Coding genetic algorithm is easy. 	<ul style="list-style-type: none"> Hard for people to come up with good heuristic Might not find optimal solution to defined problem in all cases

Conclusion

As per the comparison in Table: Comparison between machine learning techniques there are different machine learning approaches which can be used to identify the plant disease but SVM algorithm should be preferred because it gives better classification and prediction results which is helpful to identify the plant diseases.

This paper contains survey about different machine learning techniques. As per the survey, this paper has made an attempt to study machine learning method used by researchers to identify diseases and classification. These machine learning methods will help system to identify disease occurred on plant by image processing and system will inform farmer about disease in detail and specify the medicine to get rid of plant disease and increase the productivity of

References

- [1]. F Ahmed, HA AI-Mamun, ASMH Bari, E Hossain, "Classification of crops and weeds from digital images: A SVM approach", *Elsevier-2012*