A Novel Approach for the Detection of Plant Diseases

Sujeet Varshney
Research Scholar
CBS Group of Institution, Jhajjar, Haryana

Tarun Dalal
Assistant Professor (CSE Department)
CBS Group of Institution, Jhajjar, Haryana

Abstract—This proposal is about automatic detection of diseases and diseased part present in the leaf images of plants and even in the agriculture Crop production. It is done with advancement of computer technology which helps in farming to increase the production. Mainly there is problem of detection accuracy and in neural network approach support vector machine (SVM) is already exist. In this research proposal, we have discussed the various advantages and disadvantage of the plant diseases prediction techniques and proposed a novel approach for the detection algorithm, a framework of our proposed work is given in this proposal and methodology is included.

Keywords—SVM, ANN, PNN, PLANT DISEASES

I. INTRODUCTION

Agriculture has become much more than simply a means to feed ever growing populations. It is very important where in more than 70% population depends on agriculture in India. That means it feeds great number of people. Farmers are called “the backbone of India”. Plants become an important source of energy and only a primary source to the problem of global warming. The damage caused by emerging, re-emerging and endemic pathogens, is important in plant systems and leads to potential loss economically. In addition, crop diseases contribute directly and indirectly to the spread of human infectious diseases and environmental damage. As these diseases are spreading worldwide causing damage to the normal functioning of the plant and also damaging the financial condition by significantly reducing the quantity of crops grown. The crop production losses its quality due to much type diseases and sometimes they occur but are even not visible with naked eyes. Farmers estimate the diseases by their experience but this is not proper way.

Figure 1. Infected leaf of cotton plant
1.1. Plant diseases analysis and its symptoms

Detection of plant disease and assessment of the amount on individual plants or in plant populations is required where crop loss must be related to disease, for plant disease surveys, in plant breeding to assess host susceptibility, to make cost-effective disease management decisions in crop production and to better understand many basic biological processes (e.g. co-evolution). Disease assessment is also required for aiding in the settlement of crop insurance claims, aspects of crop biosecurity (biocrimes) and possibly terrorism.

The RGB image feature pixel counting techniques is extensively applied to agricultural science. Image analysis can be applied for the following purposes:
1. To detect plant leaf, stem, and fruit diseases.
2. To quantify affected area by disease.
3. To find the boundaries of the affected area.
4. To determine the color of the affected area.
5. To determine size & shape of fruits.

Following are some common symptoms of fungal, bacterial and viral plant leaf diseases.

1.1.1. Bacterial disease symptoms

Bacterial diseases include any type of illness caused by bacteria. Bacteria are a type of microorganism, which are tiny forms of life that can only be seen with a microscope. Other types of microorganisms include viruses, some fungi, and some parasites. Millions of bacteria normally live on the skin, in the intestines, and on the genitalia. The vast majority of bacteria do not cause disease, and many bacteria are actually helpful and even necessary for good health. These bacteria are sometimes referred to as “good bacteria” or “healthy bacteria.”

1.1.2. Viral disease symptoms

Symptoms of viral diseases vary depending on the specific type of virus causing infection, the area of the body that is infected, the age and health history of the patient, and other factors. The symptoms of viral diseases can affect almost any area of the body or body system. Symptoms of viral diseases can include:
• Flu-like symptoms (fatigue, fever, sore throat, headache, cough, aches and pains)
• Flu-like symptoms (fatigue, fever, sore throat, headache, cough, aches and pains)
• Gastrointestinal disturbances, such as diarrhoea, nausea and vomiting
• Irritability
• Malaise (general ill feeling)
• Rash

![Figure 2. Bacterial and Viral disease on leaves](image)

1.1.3. Fungal disease symptoms

Fungi that are common in the environment often cause fungal diseases. Most fungi are not dangerous, but some types can be harmful to health. Mild fungal skin diseases can look like a rash and are very common. Fungal diseases in the lungs are often similar to other illnesses such as the flu or tuberculosis. Some fungal diseases like fungal meningitis and bloodstream infections are less common than skin and lung infections but can be deadly. These areas are covered with white to greyish on the undersides as shown in figure 3(c).
1.2. CLASSIFICATION TECHNIQUES

This section will discuss some of the popular classification techniques that are used for plant leaf classification. In plant leaf classification, leaf is classified based on its different morphological features. Some of the classification techniques used are Neural Network, Genetic Algorithm, Support Vector Machine, and Principal Component Analysis, k-Nearest Neighbor Classifier. Plant leaf disease classification has wide application in Agriculture.

1.2.1 k-Nearest Neighbor:

K-Nearest Neighbor is a simple classifier in the machine learning techniques where the classification is achieved by identifying the nearest neighbors to query examples and then make use of those neighbors for determination of the class of the query. In KNN the classification is based on the calculation of the minimum distance between the given point and other points. As a classifier the nearest neighbor does not include any training process. It is not applicable in case of large number of training examples as it is not robust to noisy data. For the plant leaf classification the Euclidean distance between the test samples and training samples is calculated. In this way it finds out similar measures and accordingly the class for test samples. A sample is classified based on the highest number of votes from the k neighbors, with the sample being assigned to the class most common amongst its k nearest neighbors. k is a positive integer, typically small. If k = 1, then the sample is simply assigned to the class of its nearest neighbor. In binary (two class) classification problems, it is helpful to choose k to be an odd number as this avoids tied votes [19][20].

1.2.2 Support Vector Machine:

Support Vector machine (SVM) is a non-linear Classifier. This is a new trend in machine learning algorithm, which is used in many pattern recognition problems, including texture classification. In SVM, the input data is non-linearly mapped to linearly separated data in some high dimensional space providing good classification performance. Fig below shows the support vector machines concept. Multiclass classification is also applicable and is basically built up by various two class SVMs to solve the problem, either by using one-versus-all or one versus-one. The winning class is then determined by the highest output function or the maximum votes respectively [18] [19] [20].

1.2.3 Artificial Neural Network (ANN):

An Artificial Neuron is basically an engineering approach of biological neuron. ANN consists of a number of nodes, called neurons. Neural networks are typically organized in layers. In neural network each neuron in hidden layer receives signals from all the neurons in the input layer. The strength of each signal and the biases are represented by weights and
constants, which are calculated through the training phase. After the inputs are weighted and added, the result is then transformed by a transfer function into the output. The transfer functions used are Sigmoid, hyperbolic tangent functions or a step. Backpropagation is a neural network learning algorithm (Rumelhart and McClelland, 1986) used in layered feed-forward Artificial Neural Networks. Backpropagation is a form of supervised training [21][17].

![Multilayered Artificial Neural Network](image)

1.2.4 Self-Organizing Map

A self-organizing map (SOM) or self-organising feature map (SOFM) is a type of artificial neural network (ANN) that is trained using unsupervised learning to produce a low-dimensional, discretized representation of the input space of the training samples, called a map. Self-organizing maps are different from other artificial neural networks as they apply competitive learning as opposed to error-correction learning, and in the sense that they use a neighborhood function to preserve the topological properties of the input space.

A self-organizing map showing U.S. Congress voting patterns visualized in Synapse. The first two boxes show clustering and distances while the remaining ones show the component planes. Red means a yes vote while blue means a no vote in the component planes.

1.2.5 Probabilistic Neural Network

Probabilistic Neural Networks (PNNs) is a feed forward neural network, based on Parzen windows. In a PNN, the operations are organized into a multilayered feed forward network with four layers. PNN is mainly used in classification problems. The first layer is input layer, which calculates the distance from the input vector to the training input vectors. The second layer sums the contribution for each class of inputs and produces its net output as a vector of probabilities.

In a PNN, the operations are organized into a multilayered feed forward network with four layers:

- Input layer
- Pattern layer
- Summation layer
- Output layer

1.2.6 Fuzzy Logic

Fuzzy Logic classifiers are classification systems that make use of fuzzy sets or fuzzy logic (Kuncheva, 2000) which convert real-world data values into membership degrees through the use of the membership functions (Zadeh, 1965) so that these rules then can be used for the classification process. This is done by defining “categories” for each one of the attributes [18]. As Fuzzy logic classifier’s has very high speed they are preferable in cases where there is limited precision in the data values or when classification is required in real time.

II. LITERATURE REVIEW

[1] The detection of disease has been done in three steps. It uses k clustering to identify the object or the pattern of diseased image of leaf by using this technique. Then, feature extraction will be applied by the methodology named co-occurrence of texture analysis. At last detection of infected and uninfected part is done and then the classifier will be applied to it.

[2] This paper, use pattern recognition and explains how the Accuracy of plant diseases detection is to be improved. The image was taken and image recognition was done. The steps were needed to apply on it that is image compression, image cropping and denoising. Here two diseases of grapes that is grape downy mildew and grape powdery mildew and two diseases of wheat that is wheat stripe rust and wheat lead rust were detected.
[3] In this paper, the combination of two methods is used to detect the infected part of disease. Firstly, segmentation is done which is based on edge detection. It takes the image which is RGB model as input image and then feature values will be calculated and get required output. Then image analysis will be performed and after that classifier will be applied on it.

[4] This paper has mainly three objectives; first one is to eliminate both the infected part and the non-infected part from each other in the image. Secondly, to tell about the differentiation between the diseases like leaf rust Cercospora leaf spot and powdery mildew. Thirdly, to identify diseases before its symptoms started appearing or are getting visible with naked eye. First important thing in this paper was automatic detection of disease at very early stage to protect the production of the crop.

[5] Automatic detection of disease have first phase and this is that one and is very important also. A method using digital image processing technique for detecting disease like spot on plant and its segmentation and will be implanted to detect disease. In this paper a comparison of the effect of CIELAB, HSI and YCbCr colour space in the process of disease spot detection is done. Spots are basically of different colors but have not the same type of intensity with comparison to some other colours of plant leaves. As in other transformation of colour is to done, RGB model to other model for purpose of smoothing the filter named median filter is used. And then threshold can be calculated by apply some type of method to detect spot diseases also. This is very robust and efficient method. An algorithm which is independent of background noise, plant type and disease spot colour was developed and experiments were carried out on different “Monocot” and “Dicot” family plant leaves with both, noise free (white) and noisy background.

[6] In this paper the SVM has been used for detection of feature extraction and it is supervised machine learning algorithm and also nonlinear in nature and set of the related features is also extracted. In this Support Vector Machine is basically increase or improve the accuracy of detection.

[7] Recommended a k-means clustering technique for segmentation. RGB has been converted to HIS, where H is the hue, I indicate the intensity and S indicate the saturation value. Color Co occurrence method or CCM method has been used for color feature extraction. Plant disease is detected using Histogram matching. The Threshold value for the pixel is computed using Otsu’s method.

[8] proposed Rice Disease Identification using Pattern Recognition Techniques describes a software prototype system for rice disease detection based on the infected images of various rice plants. Using digital camera images of infected rice plants are captured and using image growing, image segmentation techniques to detect infected parts of the plants. Then the classification of infected part of leaf is done by neural network.

[9] the proposed methods follow three steps: Pre processing of image, Segmentation of image and Statistical analysis. First image is acquired and then color image vector median filter is used to pre-process the image of crop leaves. The processed image of crop disease is selected on the basis of affected and normal area color difference. Then normal and affected areas are segmented by statistical pattern recognition. Then ratio percentage of the normal and affected area is calculated. Finally the degree of severity of crop disease is found out. During preprocessing of image color information method is used instead of traditional image enhancement method. With the use of improved vector median filter better edges are achieved. To segment the image the statistical pattern recognition classifier was used. By the use of this classifier the crop disease images of corn, grape and cucumber were separated from the background. Good segmentation was achieved because the area can be divided by the normal area. To reduce the noise from that identification open and close mathematical operation were performed. The accuracy of grading was improved while time and cost were reduced. Hardware requirement was very low.

[10] Plant disease detection and severity assessment are required for many purposes, including predicting yield loss, monitoring and forecasting epidemics, judging host resistance and for studying fundamental biological host–pathogen processes. If assessments of disease severity are inaccurate and/or imprecise, incorrect conclusions might be drawn and incorrect actions taken. Image analysis based on digital images made using visible wavelengths is one of the several methods used to detect and quantify disease; it offers advantages compared with visual assessment or other methods. Over the last 30 years, major advances have been made to improve reliability, precision and accuracy of image analysis for detecting and measuring plant disease. Although the equipment and software continue to become more sophisticated, these technologies are also becoming easier to use. As a result, image analysis to measure plant disease is becoming increasingly widely used, and has now been applied in the study of numerous plant diseases. This review describes the history, technology and application of visible-wavelength photography and image analysis, and progress towards realizing the full potential of these systems in plant disease detection and assessment.

### III. Experimental Results

Our experimental study and its results mainly will focus on SVM algorithm. This algorithm will first take an input image, which is in RGB form. It detects the infected part of the disease. We have taken two data sets mainly named train dataset and training data set. Train data set contains images, which already have been processed, and diseases and feature extraction about them have been detected.
Another dataset named training dataset contain the images, which needed to be processed for detection of diseases. One image from training data set is taken and its features are matched to images in the training dataset. Main thing in this work is that whole area, which is infected by disease, will be calculated in the terms of percentage %. And disease is also detected. Now following are the results of the implemented work.

GUI of our proposed work is shown in Fig 4.1. After clicking on Load Image button, this button load an image manually from system and after loading an image the image is displayed in the query image box below the load image. This action load the image on GUI which is selected by user and the pictorial view is given in the below Fig 4.2.

![Fig 4.1 Main GUI of proposed work](image)

After loading an image, next work is the enhancing the contrast from the image. SVM enhance the image very accurately. This action is performed by clicking on the Enhance Contract button next to the Load Image. After click on this button all the features are enhanced from the loaded image and the new image is generated in contrast enhanced box. The results is shown in the below Fig 4.3.

![Fig 4.2 Loading of an Image](image)

To wok on our area of interest we have to segment the image. After clicking on the Segment image button, we got an image, which is segmented in our region of interest (ROI). Below Fig 4.4 shows the segmented ROI results.

![Fig 4.3 Contrast Enhanced](image)
Classification results in the below Fig 4.5 shows that the loaded image of the leaf is the healthy leaf and there is no affected regions in the leaf. Corresponding features of the leaf image is calculated and results are displayed in Fig 4.5.

Accuracy of the our proposed results is 96.77% as shown in the below Fig 4.6. This results proof that our proposed algorithm is much accurate as compare to existing one.

After calculating the accuracy of the healthy leaf, we upload a new image for finding the affected region and the accuracy of the proposed algorithm. A new image loaded is shown in Fig 4.7.
Fig 4.7 Loading of an image

Same as the first image, the contrast enhancing is performed on the image shown in below Fig: 4.8.

Fig 4.8 Contrast Enhancing of 2nd image

Segmenting the interested region from the Image shows the results as given in 4.9. As it is clear from the Fig. that segmentation result is very dark as compare to loaded image. After segmentation corresponding features of the leaf image is calculated and results are displayed in Fig 4.9.

Fig 4.9 Segment ROI of second image

Classification results in the below Fig 4.10 shows that the loaded image of the leaf is the disease affected leaf and name of the disease is Cercospora leaf spot. There is 30.439% affected regions in the leaf. Corresponding features of the affected leaf image is calculated and results are displayed in Fig 4.10.
Accuracy of the our proposed results of the affected leaf is 98.38% as shown in the below Fig 4.11. This results proof that our proposed algorithm is much accurate as compare to existing one.

Below table and bar graph comparison shows the comparison of accuracy on four different images.

Table Comparison

<table>
<thead>
<tr>
<th></th>
<th>Existing Method</th>
<th>Proposed Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Image 1</td>
<td>79%</td>
<td>98.3871%</td>
</tr>
<tr>
<td>Image 2</td>
<td>76%</td>
<td>96.7742%</td>
</tr>
<tr>
<td>Image 3</td>
<td>70%</td>
<td>90.4312%</td>
</tr>
<tr>
<td>Image 4</td>
<td>69%</td>
<td>88.6641%</td>
</tr>
</tbody>
</table>
IV. CONCLUSIONS
The proposal reviews and summarizes some techniques that have been used for plant disease detection. The two major classification techniques of plant diseases. The spectroscopic and imaging techniques include fluorescence spectroscopy, visible-IR spectroscopy, fluorescence imaging, and hyper spectral imaging. A novel approach for disease prediction of plants based on classification technique is proposed.

Acknowledgement
Every success stands as a testimony not only to the hardship but also to hearts behind it. Likewise, the present work has been undertaken and completed with direct and indirect help from many people and I would like to acknowledge all of them for the same.

REFERENCES