# Plant Disease Detection using Image Processing- A Review

Surender Kumar
Department of CSE
Chandigarh University
Gharuan, Punjab

Rupinder Kaur
Assistant Professor
Department of CSE
Chandigarh University
Gharuan, Punjab

## ABSTRACT
This paper holds a survey on plant leaf diseases classification using image processing. Digital image processing has three basic steps: image processing, analysis and understanding. Image processing contains the preprocessing of the plant leaf as segmentation, color extraction, diseases specific data extraction and filtration of images. Image analysis generally deals with the classification of diseases. Plant leaf can be classified based on their morphological features with the help of various classification techniques such as PCA, SVM, and Neural Network. These classifications can be defined various properties of the plant leaf such as color, intensity, dimensions. Back propagation is most commonly used neural network. It has many learning, training, transfer functions which is used to construct various BP networks. Characteristics features are the performance parameter for image recognition. BP networks shows very good results in classification of the grapes leaf diseases. This paper provides an overview on different image processing techniques along with BP Networks used in leaf disease classification.

## Keywords
Back Propagation; Image Processing; Artificial Neural Network

## 1. INTRODUCTION
Since the ancient times agriculture has its own importance in human life. Plants are the basic source for supply of energy for human body. Productions based on agriculture get easily affected by various plant diseases. These diseases cost as social, ecological and economical loss to farmers. It becomes very important to analyze plant diseases very accurately within specific time. Some diseases are visible to human eyes and can be easily detected and procured. Some are so sophisticated needs powerful microscopes or specific electromagnetic spectrum. Digital technology can make it very easy task to process all kind of disease images very precisely. It also gives the facility to remote sense the diseases without having an expert on the field. Fungi diseases can be detected by implementation of various automated algorithms of neural networks e.g. back propagation, PCA, etc.

Back propagation is an approach to train undefined neural networks. In this mathematical algorithm, errors from output propagate to the input for adjustment of weights. Any number of hidden unit layers can be trained with back propagation algorithms. If vector x and t represents input pattern and desired target output respectively, $w_{ij}$ represents weight between node i and node j.

## 2. DEFINITIONS
The important terms are defined as below:

### 2.1 Error Signal and Weight gradient
Error signal δ at a node j is:

$$\delta_j = -\partial E / \partial net_j$$

The weight $w_{ijv}$ gradient is given by:

$$\Delta w_{ij} = -\partial E / \partial w_{ij}$$

The set of nodes anterior to node i and node j:

$$A_i = \{ j : \exists w_{ij} \}$$
$$P_j = \{ i : \exists w_{ij} \}$$

### 2.2 The gradient
Use chain rule to expand gradient. It creates two factors

$$\Delta w_{ij} = -\frac{\partial E}{\partial net_i} \frac{\partial net_i}{\partial w_{ij}}$$

First factor contains error of node i. The second is

$$\frac{\partial net_i}{\partial w_{ij}} = \frac{\partial}{\partial w_{ij}} \sum_{k \in A_i} w_{ik} \, y_k = y_j$$

When put two equations together

$$\Delta w_{ij} = \delta_i \, y_j.$$

Gradient can be calculated when all the activity and error relevant to nodes in network is known.

### 2.3 Forward activation
Input units activity depends upon external input x, for rests of others previous activity propagates.

$$y_i = f_i \left( \sum_{j \in A_i} w_{ij} \, y_j \right)$$

To calculate a activity of particular nodes, activity for all its ancestors should to known. Nodes order from input to output should be maintained as feed forward network.

### 2.4 Output error
Using sum-squared loss

$$E = \frac{1}{2} \sum_{O} (t_O - y_O)^2$$

Error for output unit o would be:

$$\delta_O = t_O - y_O$$

## 2.5 Error Back propagation

To train hidden node, error should propagate back from output node to input node. Again expand hidden unit error using chain rule for posterior nodes:

$$\delta_j = -\sum_{i \in P_j} \frac{\partial E}{\partial net_i} \frac{\partial net_i}{\partial y_j} \frac{\partial y_j}{\partial net_j}$$

The first term represents error of node i. The second term is given be:

$$\frac{\partial net_i}{\partial y_j} = \frac{\partial}{\partial y_j} \sum_{k \in A_i} w_{ik}\ y_k = w_{ij}$$

The third derivative shows activation function of node j:

$$\frac{\partial y_j}{\partial net_j} = \frac{\partial f_j(net_j)}{\partial net_j} = f_j^{'}(net_j)$$

tanh activation function can be used for hidden unit h using special identity

$$\tanh(u)' = 1 - \tanh(u)^2,$$

This gives

$$f_j^{'}(net_h) = 1 - y_h^2$$

Putting all the pieces together, we get

$$\delta_j = f_j^{'}(net_j) \sum_{i \in P_j} \delta_i\ w_{ij}$$

Errors of any unit only can be calculated if all the posterior node errors are already known. This note is applicable in both directions.

## 3. RELATED WORK

Savita N. et al. [1] studied various classification techniques for plant leaf diseases. Each pattern of distinct classes is classified in classification technique based on their morphological features. Various techniques such as Artificial neural network, Probabilistic Neural Network, Genetic Algorithm, k-Nearest Neighbor, Principal Component Analysis and Fuzzy logic. It is a difficult task to select a classification method because every classification methods have its own disadvantage and advantage. This paper helps to finalize which classification method is suitable for a particular application. K-Nearest-Neighbor method is one of the simplest algorithms to test classes but it is very time complex with making predictions. Neural networks have ability to tolerate noisy input but having hard to understand algorithm structure. For classifying high-dimensional data set SVM (Support Vector Machine) found to be the best available machine learning algorithm. In SVM, quadratic optimization helps to control complexity of frequency of error and decision rule but it is very difficult to find optimal parameter for training non linear data which makes SVM more complex to implement.

Prof. Sanjay et al. [2] introduced that automatic detection of plant leaf can prevent serious outbreak. The fact that morphology with their reproductive structure can be used to identify fungi was identified. Bacteria have simpler life cycle and generally have single cell. It divides into two cells during binary fission and increase in numbers. Viruses are non protein associated protein and genetic material based extremely small particle. They developed a processing scheme with four steps. First, takes input RGB image and creates color transformation. Second, Conversion from RGB to HIS take place for color descriptor. Third, masking of pixel and removal of unwanted color using specific threshold, based on

that extraction of useful segmentation has been done. Fourth, compute the statistics from SGDM matrices and evaluate plant leaf diseases.

Mr. Pramod and S. landgeet [3] proposed an image processing based software methodology for plant diseases detection and classification. They explains how Indian farmers faces problem with plant diseases due to lack of agricultural expert access. In this paper, their objective was to develop a image processing technique to automatically recognize plant diseases based on their shape, texture and color. After detection of cause or plant diseases this system gives fast and instant information to farmers with SMS. This system will reduce cost, chemical testing procedure, time and enhance productivity.

Anand. H. Kulkarni and Ashwin Patil R. K [4] described a methodology to accurately detecting plant diseases. This methodology includes image processing techniques along with artificial neural network (ANN). They also discuss about the dramatic problems of farmers and present a work aimed to develop diseases detection system for plant. This system first capture raw image of plants. Then, Filter and Segment it by Gabor filter. Then, extract the color information from segmented image. Now well trained ANN is used to distinguish between healthy and diseased plant sample. Color and texture are two precious parameters for ANN based classifier. This system runs with real time constraints and shows 91% accuracy on implementation.

Haiguang Wang et al. [5] proposed a model to improve accuracy of image recognition and diagnosis of plant diseases. Two kinds of wheat disease (wheat stripe rust and leaf rust) and grapes diseases (grape downy mildew and powdery mildew) were under inspection. Image processing which includes image compression, image cropping and image denoising along with k-means clustering algorithms for segmentation had been used for image recognition. 25 texture features, 21 color features and 4 shape features were extracted from the each image. Backpropagation (BP) algorithms based classifier used to detect diseases. Results show, backpropogation networks are extremely effective to identification of diseases. Principal Component Analysis (PCA) is used to features the dimension data. As a optimal resolution result prediction accuracy and fitting accuracy for the grape diseases were 100%. For wheat too it is 100% both. When feature data dimensions were reduced by PCA, prediction accuracy was 97.14% and fitting accuracy was 100% for grapes diseases optical recognition, andboth fitting accuracy and the prediction accuracy were 100% for wheat diseases.

Piyush Chaudhary et al. [6] proposed an algorithm for spot segmentation for plant leaf diseases using image processing technique. According to the classification of plant diseases is the very first and significant stage for plant detection. In comparison to plant leaf color, diseases spots are same in colors but different in intensities. So RBG color transform can be a better choice for diseases spot segmentation. In this paper, effect of HSI, CIELAB color and YCbCrspaces comparison has been done for spot detection. Image smoothing is achieved through median filter. Otsu method of threshold calculation is applied to detect disease spot on color component. Various "Dicot" and "Monocot" family plants leaves were analyzed in both noisy and noise free (white) background. Developed algorithm is independent of diseases spot color, plant type and background noise.

M. Egmont-Petersen [7] categorized various applications for image processing algorithms. Categorization is achieved with as a two dimensional taxonomy. One dimension of which specifies object recognition, data reduction/feature extraction, preprocessing, segmentation, optimization and image understanding. Second dimension takes input data and perform various abstraction level task as structure-level, object-set-level, pixel-level, object-level local feature-level and scene characterization.

Muhammad Faisal Zafar et al. [8] proposed the grape leaf diagnosis system. It has three main parts: firstly complex back ground color extraction, secondly color extraction diseased part of leaf and at last disease classification. Color analysis has been done with the help of self-organizing feature map with back propogation neural network. GA and MSOFM are employed for leaf disease segmentation further SVM is deployed for classification. After these processes filtration of image has been done through Gabor Wavelet and applied to SVM for classification of final disease. Grapes leaf diseases are classify into three categories in this system: Rust disease, Scab disease and No disease. This system shows wide performance for agricultural implementation.

Jayme Garcia and Arnal Barbedo [9] presented a survey on various image processing methodology to classify, quantify and detect plant diseases in a visible spectrum. The analysis has been done only on the leaves on the system to keep the survey on short. Smita Naikwadi and Niket Amoda [10], show a software evolution for plant leaf diseases detection and classification. Two steps are successively added after the segmentation process. Firstly they identify pixels having most green color. Secondly mask this pixel with Otsu's method on a specific threshold value. Addition to that, all the pixels with zeros, red, green and blue values and boundary pixels has been completely removed. Experimental results shows robust technique for plant leaf diseases is obtained.

# 4. IMAGE PROCESSING

Digital signal processing is the methodology to achieve fast and accurate result about the plant leaf diseases. it will reduce many agricultural aspect and improve productivity by detecting the appropriate diseases. For diseases detection image of an infected leaf should examine through the set of procedures. As fig 1 shows, input image should pre-processed then its feature should be extracted according to the dataset. After then some classifier techniques should be used to classify the diseases according to the specific data set.

**Image Acquisition** is the process in which acquired and converted to the desired output format. For this application an analog image is first captured and then converted to the digital image for further processing.

**Preprocessing Segmentation** contains process for image segmentation, image enhancement and color space conversion firstly image digital image is enhanced by filter. Leaf image is filtered form the background image. Then filtered image's RGB colors are converted into color space parameter. Hue Saturation Value (HSV) is a good method for color perception. Further image is segmented to a meaning full part which is easier to analyze. Any of the model based, threshold based, edge based, Region based and feature based segmentation has been done on the images.
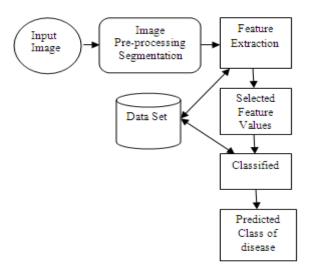


**Fig 1: Basic Steps in Image processing to detect plant diseases**

**F**eature extractions, is the process done after segmentation. According to the segmented information and predefined dataset some features of the image should be extracted. This extraction could be the any of statistical, structural, fractal or signal processing. Color co-occurrence Method, Grey Level Co-occurrence Matrices (GLCM), Spatial Gray-level Dependence Matrices (SGDM) method, Gabor Filters, Wavelets Transform and Principal component analysis are some methods used for feature extraction.

**Table 1. Classification techniques Comparison [11]**

| Technique | Advantages | Disadvantages |
|---|---|---|
| K-Nearest Neighbor (KNN) | Simpler classifier as exclusion of any training process. Applicable in case of a small not trained dataset. | More training samples- more speed of computing distances sensitive to irrelevant inputs so expensive testing every time. |
| Radial Basis Function (RBF) | Faster Training. Hidden layer is easier to interpret. | It is slower in execution speed |
| Probabilistic Neural Networks (PNN) | Tolerant to noisy inputs. Instances classified to many output adaptive to change data. | Long training time. Complex network structure. Excessive memory for training data. |
| Back-propagation Network (BPN) | Easy to implement. Applicable to wide range of problems. Able to form arbitrarily complex nonlinear mappings | Learning can be slow. It is hard to know how many neurons as well as layers are required. |
| Support Vector Machine (SVM) | Simple geometric interpretation and a sparse solution. robust, when sample has some bias. | Slow training. Difficult to understand. For classification large support vector required |

**Classifiers** are the software routine written on the platform to define certain features for classification of the images. K-nearest neighbor, Radial basis function, Probabilistic Neural Network (PNN), Convolutional neural network, Support vector machine and Back propagation network are the some linear and non linear classifier for image classification. Any platform such as MATLAB can be used to train and test these classifiers.

## 5. CONCLUSION

This paper reviews and summaries techniques used in image processing for plant leaf diseases. Very regularly general implemented techniques are SVM, BPNN, SGDM and K means Clustering. With the help from any of these techniques the automation in plant leaf disease detection is achieved. SVM is very difficult to find optimal parameters for training purpose because they are non linearly separable. Neural network can tolerate noisy inputs. K-means NN have disadvantage of time complexity while making prediction. Having these disadvantages associated with each methodology makes each individual to use for some particular reason. Which makes BPNN is the most easily implemented non linear technique for training of neural network based plant diseases detection.

## 6. REFERENCES

[1] Savita N. Ghaiwat, ParulArora, "Detection and Classification of Plant Leaf Diseases Using Image processing Techniques: A Review*", International Journal of Recent Advances in Engineering and Technology, ISSN: 2347-2812, Volume 2, Issue 3, 2014.*

[2] Prof. Sanjay B. Dhaygude and Mr. Nitin P. Kumbhar, "Agricultural plant Leaf Disease Detection Using Image Processing" *International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, Volume 2, Issue , 201*3.

[3] Mr. Pramod and S. landge , "Automatic Detection and Classification of Plant Disease through Image Processing", *International Journal of Advanced Research in Computer Science and Software Engineering, Volume 3, Issue 7, ISSN: 2277 128X, 2013.*

[4] Anand H. Kulkarni and Ashwin Patil R. K, "Applying image processing technique to detect plant diseases", *International Journal of Modern Engineering Research (IJMER), Volume 2, Issue.5, pp-3661-3664 ISSN: 2249-6645, 2012*

[5] Haiguang Wang, Guanlin Li, Zhanhong Ma, Xiaolong Li, "Image Recognition of Plant Diseases Based on Back propagation Networks", *5th International Congress on Image and Signal Processing (CISP), 2012*

[6] Piyush Chaudhary Anand K. Chaudhari, Dr. A. N. Cheeranand Sharda Godara "Color Transform Based Approach for Disease Spot Detection on Plant Leaf", *International Journal of Computer Science and Telecommunications, Volume 3, Issue 6, 2012.*

[7] M. Egmont-Petersen, " Image processing with neural networks", Elsevier, *Volume 35, Issue 10, October 2002, Pages 2279–2301 2002*

[8] Muhammad Faisal Zafar, Dzulkifli Mohamad, Muhammad Masood Anwar "Recognition of Online Isolated Handwritten Characters by Back propagation Neural Nets Using Sub-Character Primitive Features", *IEEE Multitopic Conference, 2006.*

[9] Jayme Garcia, ArnalBarbedo, "Digital image processing techniques for detecting, quantifying and classifying plant diseases", *Springer Plus, 2013.*

[10] SmitaNaikwadi, NiketAmoda, "Advances in Image processing for detection of plant disease", *International Journal of Application or Innovation in Engineering & Management, Volume 2, Issue 11, 2013.*

[11] Ms. Kiran, R. Gavhale, Prof. Ujwalla Gawande, "*An Overview of the Research on Plant Leaves Disease detection using Image Processing Technique*" *IOSR Journal of Computer Engineering (IOSR-JCE) e-ISSN: 2278-0661, p- ISSN: 2278-8727Volume 16, Issue 1, PP 10-16, 2014.*