Agricultural Crop Leaf Disease Detection

Pritam Sontakke[1], Rahul Bhor[2], Sachin Kharmale[3], Prof. Renuka Gound[4]
Department of Information Technology
[1, 2, 3, 4] Pimpri Chinchwad college of Engineering, Pune-411044

Abstract-*A plant leaf disease detection is a very important to prevent loss in agriculture. Detection of plant leaf disease by automated system is research area. Crop diseases are caused by various microorganisms like fungi, bacteria, and viruses. Conventional disease detection systems used to diagnose the disease in agriculture depends only on manual textual work. In our proposed system, we are going to develop an image processing system to automate the detection of the disease type. Early information on plant disease detection can facilitate the control of diseases by using proper strategies. Hence the algorithm is to design, implement and evaluate an image processing based software for automatic detection and classification of plant leaf diseases. The proposed system consists of four main steps, in first step color transformation structure for the input RGB image is created, in next step RGB is converted to HSI because RGB is for color generation. Specific threshold value is used to masking and removing green pixels in leaf, then segmentation of image is created and the useful segments are extracted, then the feature statistics computed. Finally detection of diseases on the plant leaf is evaluated.*

Keywords- Image acquisition, Image pre-processing, Segmentation, k-mean clustering, feature extraction, classification.

I. INTRODUCTION

As the India is an agricultural country where almost 70% of population is dependent on the agriculture. So there is need for optimum and quality crop to be yield. It can be improved by the aid of technological support. But, now a day's people are following the general method for detection and identification of the plant disease that is naked eye observations [1]. But, this requires the continuous monitoring of the experts which is truly expensive in the large farms. Automatic detection of the plant disease is an essential research topic as they can improve the monitoring of large farms and thus automatically detects the symptoms of the disease that appear on the plant leaves. Therefore looking for fast, automatic, less expensive and accurate method to calculate plant disease by calculating plant area through pixel statistics [1]. The Image processing can be used for agricultural purpose as follow:

- To detect diseased leaf, stem, fruit.
- To quantify affected area by disease.
- To find shape of affected area.
- To find the patterns and color of the area affected. Etc.

The Image processing technique is used for preprocessing, features are extracted and classified by any of classification techniques like SVM classifiers, decision tree. Automatic classification of leaf disease is done based on the high resolution multispectral and stereo images. Segmentation is color based segmentation used to obtain the diseased region of leaf [2].

II. LITERATURE SURVEY

In this section, various method of image processing for plant disease detection is discussed. The vegetation indices from hyper spectral data have been shown for indirect monitoring of plant diseases. But they cannot distinguish different diseases on crop. Wenjiang Huang winter wheat disease. They consider three different pests (Powdery mildew, yellow rust and aphids) in winter wheat for their study. The most and the least relevant wavelengths for different diseases were extracted using RELIEF-F algorithm. The classification accuracies of these new indices for healthy and infected leaves with powdery mildew, yellow rust and aphids were 86.5%, 85.2%, 91.6% and 93.5% respectively [9].

Enhanced images have high quality and clarity than the original image. Color images have primary colors red, green and blue. It is difficult to implement the applications using RGB because of their range i.e. 0 to 255. Hence they convert the RGB images into the grey images. Then the histogram equalization which distributes the intensities of the images is applied on the image to enhance the plant disease images. Monica Jhuria et al uses image processing for detection of disease and the fruit grading in [10]. They have used artificial neural network for detection of disease. They have created two separate databases, one for the training of already stored disease images and other for the implementation of the query images. Back propagation is used for the weight adjustment of training databases. They consider three feature vectors, namely, color, textures and morphology [10]. They have found that the morphological feature gives better result than the other t wo features. Zulkifli Bin Husin et al, in their paper [11], they captured the chilli plant leaf image and processed to determine the health status of the chilli plant. Their technique is ensuring that the chemicals should apply to

the diseased chilli plant only. They used the MATLAB for the feature extraction and image recognition. In this paper pre-processing is done by converting image into HIS image and then contrast is enhanced. The segmentation of leaf image is important while extracting the feature from that image. Mrunalini R. Badnakhe, Prashant R. Deshmukh compare the Otsu threshold and the k-means clustering algorithm used for infected leaf analysis in [5]. They have concluded that the extracted values of the features are less for k-means clustering. The clarity of k-means clustering is more accurate than other method.

The RGB image is used for the identification of disease. After applying k-means clustering techniques, the green pixels is identified using color based segmentation. For the feature extraction, color co-occurrence method is used. RGB image is converted into the HSI translation. For the texture statistics using GLCM function the feature is calculated [12]. The FPGA and DSP based system is developed by Chunxia Zhang, Xiuqing Wang and Xudong Li, for monitoring and control of plant diseases [13]. The FPGA is used to get the fieldplant image or video data for monitoring and diagnosis. The DSP TMS320DM642 is used to process and encode the video or image data. The nRF24L01 single chip 2.4 GHz radio transmitter is used for data transfer. It has two data compress and transmission method to meet user's different need and uses multi-channel wireless communication to lower the whole system cost. Shantanu Phadikar and Jaya Sil uses pattern recognition techniques for the identification of rice disease in [9]. This paper describes a software prototype for common leaf disease detection based on infected image of plant.

III. THE PROPOSED APPROACH

Image processing is the technique for enhancing the image so that results are more suitable for a particular application. In this section the basic steps for disease identification and classification are shown (Fig. 1).

A] Image Acquisition

The images of the plant are captured through the camera and these images are in RGB format. The color transformation structure is then created for RGB image. Then the device-independent color space transformation structure is applied.

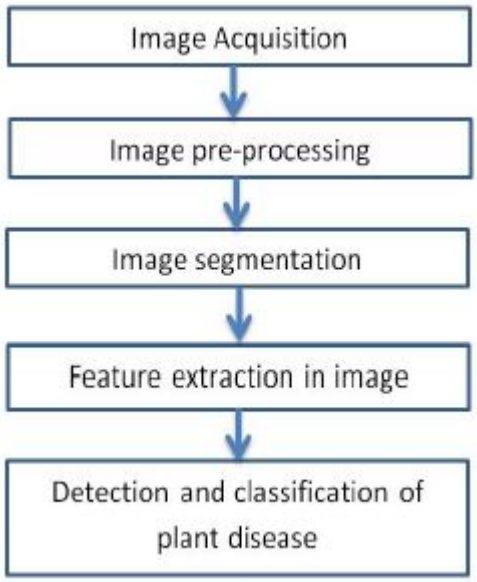

Fig. 1 Basic Steps for plant disease detection and identification [1]

B] Image Preprocessing

To remove the noise as well as to extract the objects from the image different preprocessing techniques are used. Such as, Image clipping or Image cropping to get the required area only. Image smoothening is done using smoothening filter and image contrast is enhanced [1]. The RGB images into the grey images using color conversion using equation (1).

$$f(x)=0.2989*R + 0.5870*G + 0.114.*B \text{ - - - - - -} \quad (1)$$

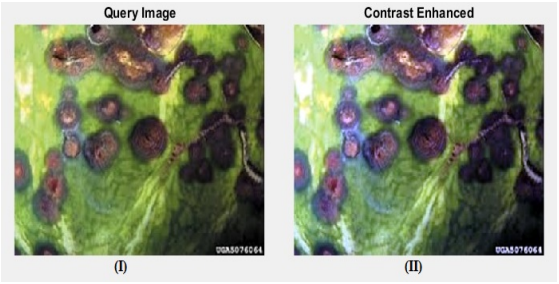

Fig. 2 Image Preprocessing

C] Image Segmentation

Segmentation is the process of segmenting the target leaf image into various part of features ore they having some similarity that means color based segmentation. The segmentation is carried out using some conventional methods such as K-means clustering, converting RGB model to HSI model etc.

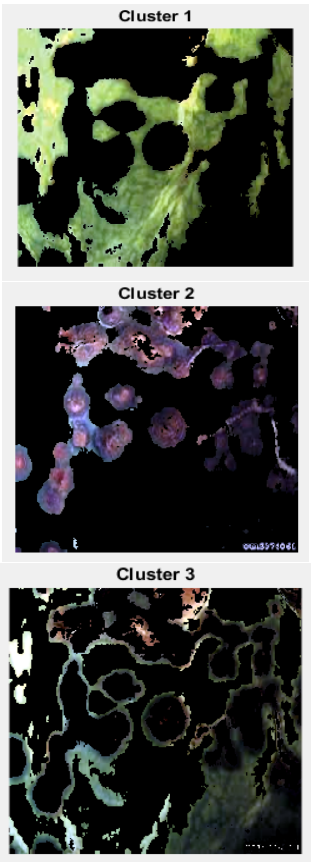

Fig. 3 Segmentation and Clustering

Image Segmentation algorithms:

1] Otsu Threshold Algorithm:

Thresholding creates binary images from grey-level images by setting all pixels below some threshold to zero and all pixels above that threshold to one. The Otsu algorithm defined in [5] is as follows:

i) According to the threshold, separate pixels into two clusters
ii) Then find the mean of each cluster.
iii) Square the difference between the means.
iv) Multiply the number of pixels in one cluster times the number in the other

The infected leaf shows the symptoms of the disease by changing the color of the leaf. Hence the greenness of the leaves can be used for the detection of the infected portion of the leaf. The R, G and B component are extracted from the image. The threshold is calculated using the Otsu's method. Then the green pixels is masked and removed if the green pixel intensities are less than the computed threshold.

2] K-means clustering algorithm:

The K-means clustering algorithm is used for classification of image objects based on a set of features into K number of classes. The classification of object is done by minimizing the sum of the squares of the distance between the object and the corresponding cluster [2].

The algorithm for K –means Clustering:

1. Decide the center of K cluster, either randomly or based on some heuristic method.
2. Assign each pixel in the image to the cluster that minimizes the distance between the pixel and the cluster center.
3. Again compute the cluster centers by averaging all of the pixels in the cluster. Repeat steps 2 and 3 until convergence is attained.

D] Feature Extraction

Using GLCM (Gray Level Co-occurrence Matrices) method features are extracted from the diseased leaf image. Color, texture, morphology and edges are some of the features can be extracted by this method and these features are used for disease classification [1].

Color Co-occurrence method:

Here, in this method both the color and textures are considered to get unique features of the image. For this RGB is converted into HSI model.

$$H(\text{Hue}) = \cos^{-1}[(R - \tfrac{1}{2}G - \tfrac{1}{2}B)/\sqrt{R^2 + G^2 + B^2 - RG - RB - GB}] \quad \text{if } G \geq B, \quad [6]$$

or

$$H(\text{Hue}) = 360 - \cos^{-1}[(R - \tfrac{1}{2}G - \tfrac{1}{2}B)/\sqrt{R^2 + G^2 + B^2 - RG - RB - GB}] \quad \text{if } B > G. \quad [6]$$

$$S(\text{Saturation}) = 1 - m / I \quad \text{if } I > 0,$$

or

$$S = 0 \quad \text{if } I = 0. \quad [6]$$

$$I(\text{Intensity}) = (R + G + B)/3. \quad [6]$$

E] Classification

Classification is final stage of this project which actually detects disease class of plant leaf. Classification is done by different classification techniques such as SVM classifier, and Decision tree algorithm etc. In this project we have classified using ID3(Iterative Dichotomiser 3) decision tree algorithm.

ID3 (Iterative Dichotomiser 3) algorithm:

In decision tree learning, ID3 (Iterative Dichotomiser

3) is an algorithm invented by Ross Quinlan used to generate a decision tree from a dataset[8].

Working of ID3:

1. Calculate the entropy of every attribute using the data set.
2. Calculate the entropy of every attribute using the data set
3. Calculate the entropy of every attribute using the data set
4. Recurse on subsets using remaining attributes[8].

The ID3 metrics:

Entropy:

Entropy H(S) is a measure of the amount of uncertainty in the (data) set S (i.e. entropy characterizes the (data) set S) [8].

$$H(S) = \sum_{x \in X} p(x) \log_2 (1/p(x))$$

Where,

- S – The current (data) set for which entropy is being calculated (changes every iteration of the ID3 algorithm)
- X – Set of classes in S
- p(x) – The proportion of the number of elements in class x to the number of elements in set S
- When H(S)=0, the set S is perfectly classified (i.e. all elements in S are of the same class) [8].

Information gain:

Information gain IG(A) is the measure of the difference in entropy from before to after the set S is split on an attribute A. In other words, how much uncertainty in S was reduced after splitting set S on attribute A[8].

$$IG(A, S) = H(S) - \sum_{t \in T} p(t) H(t)$$

Where,

- H(S) – Entropy of set S
- T – The subsets created from splitting set S by attribute A such that S=$\bigcup_{t \in T} t$
- p(t) – The proportion of the number of elements in t to the number of elements in set S
- H(t) – Entropy of subset t[8].

In classification part we have calculated global entropy of dataset which are texture features values, entropy of each feature to calculate information gain of each feature. Finally ID3 is applied to generate decision tree and target leaf disease is classified using tree.

IV. CONCLUSION

The correct detection and classification of the plant disease is very much important for efficient crop cultivation and this is possible because of image processing techniques. This paper discussed various techniques to segment the disease part of the plant leaf. This paper also discussed some Feature extraction and classification techniques to extract the features of infected leaf and the classification of plant diseases. The use of K-means clustering for cluster analysis and Multiclass SVM for classification of disease in plants. From these methods, we can accurately identify and classify various plant diseases using image processing techniques.

REFERENCES

[1] Sachin D. Khirade, A.B. Patil, Plant Disease Detection Using Image Processing, 2015, International Conference on Computing Communication Control and Automation

[2] Prof. Sanjay B. Dhaygude, Mr. Nitin P. Kumbhar, Agricultural Plant leaf Disease detection using Image Processing, International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering Vol. 2, Issue 1, January 2013 ISSN: 2278-8875

[3] Vidyaraj K, Priya S, Developing an algorithm for Tomato leaf disease detection and classification, International journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering, Vol. 3 Spacial issue 1, February 2016 ISSN 2321-2014

[4] Dr.K.Thangadurai, K.Padmavathi, "Computer Visionimage Enhancement For Plant Leaves Disease Detection", 2014 World Congress on Computing and Communication Technologies.

[5] Mrunalini R. Badnakhe, Prashant R. Deshmukh, "Infected Leaf Analysis and Comparison by Otsu Threshold and k-Means Clustering", International Journal of Advanced Research in Computer Science and Software Engineering, Volume 2, Issue 3, March 2012.

[6] http://www.had2know.com/technology/hsi-rgb-color-converter-equations.html

[7] http://www.eng.cam.ac.uk/~kkc21/thesis_main/node23.html

[8] https://en.wikipedia.org/wiki/ID3_algorithm

[9] Wenjiang Huang, Qingsong Guan, Juhua Luo, Jingcheng Zhang, Jinling Zhao, Dong Liang, Linsheng Huang, and Dongyan Zhang, "New Optimized Spectral Indices for Identifying and monitoring Winter Wheat Diseases", IEEE journal of selected topics in applied earth observation and remote sensing,Vol. 7, No. 6, June 2014

[10] Monica Jhuria, Ashwani Kumar, and Rushikesh Borse, "Image Processing For Smart Farming: Detection Of Disease And Fruit Grading", Proceedings of the 2013

IEEE Second International Conference on Image Information Processing (ICIIP-2013

[11] Zulkifli Bin Husin, Abdul Hallis Bin Abdul Aziz, Ali Yeon Bin Md Shakaff Rohani Binti S Mohamed Farook, "Feasibility Study on Plant Chili Disease Detection Using Image Processing Techniques", 2012 Third International Conference on Intelligent Systems Modelling and Simulation.

[12] H. Al-Hiary, S. Bani-Ahmad, M. Reyalat, M. Braik and Z. ALRahamneh, "Fast and Accurate Detection and Classification of Plant Diseases", International Journal of Computer Applications (0975 – 8887)Volume 17– No.1, March 2011

[13] Chunxia Zhang, Xiuqing Wang, Xudong Li, "Design of Monitoring and Control Plant Disease System Based on DSP&FPGA", 2010 Second International Conference on Networks Security, Wireless Communications and Trusted Computing.