IJCSE International Journal of Computer Sciences and Engineering Open Access
Survey Paper Vol.-7, Issue-2, Feb 2019 E-ISSN: 2347-2693

Diseases Identification in Plants Using K-Means Algorithm

M.Krishnamoorthy[1*], A. Noble Mary Juliet[2], C. Keerthana[3], R.Usha Nandhini[4]

[1,2,3,4] Department of Computer Science and Engineering, Dr. Mahalingam College of Engineering and techonology, Pollachi, Tamilnadu

[*]*Corresponding Author: krishnamurugavel3028@gmail.com, Tel.: +91 95666 07934*

DOI: https://doi.org/10.26438/ijcse/v7i2.458462 | Available online at: www.ijcseonline.org

Accepted: 19/Feb/2019, Published: 28/Feb/2019

***Abstract*—** The detection of leaf is a very important factor to prevent serious outbreak. Automatic detection of plant disease is essential. Most plant diseases are caused by fungi, bacteria, and viruses. Fungi are identified primarily from their morphology, with emphasis placed on their reproductive structures. Bacteria are considered more primitive than fungi and generally have simpler life cycles. With few exceptions, bacteria exist as single cells and increase in numbers by dividing into two cells during a process called binary fission. Viruses are extremely tiny particles consisting of protein and genetic material with no associated protein. The term disease is usually used only for the destruction of live plants. The developed processing scheme consists of four main steps, first a color transformation structure for the input RGB image is created, and this RGB is converted to HSI because RGB is for color generation and for color descriptor. Then green pixels are masked and removed using specific threshold value, then the image is segmented and the useful segments are extracted. Finally the presence of diseases on the plant leaf is evaluated.

***Keywords*—**Noise removal, Segmentation, clustering, pre-processing.

I. INTRODUCTION

This India is fast developing country and agriculture is the back bone for the development in the early stages. Due to industrialization and globalization concepts the field is facing hurdles. On top of that the awareness and the necessity of the cultivation need to be in still in the minds of the younger generation. Now a day's technology plays vital role in all the fields but till today we are using some old methodologies in agriculture. Identifying plant disease wrongly leads to huge loss of yield, time, money and quality of product. Identifying the condition of plant plays an important role for successful cultivation. In olden days identification is done manually by the experienced people but due to the environment changes the prediction is becoming tough. So we can use image processing techniques for identification of plant disease. Generally we can observe the symptom of disease on leafs, stems, flowers etc. So here we use leafs for identification of disease affected plants.

Digital image process is the use of computer algorithms to perform image process on digital pictures. It permits a far wider vary of algorithms to be applied to the computer file and might avoid issues like the build-up of noise and signal distortion throughout process. Digital image process has terribly important role in agriculture field. It's widely accustomed observe the crop disease with high accuracy. Detection and recognition of diseases in plants mistreatment digital image method is extremely effective in providing symptoms of characteristic diseases at its early stages. Plant pathologists will analyze the digital pictures mistreatment digital image process for diagnosing of crop diseases. Computer Systems area unit developed for agricultural applications, like detection of leaf diseases, fruits diseases etc. Altogether these techniques, digital pictures are collected employing a camera and image process techniques are applied on these pictures to extract valuable data that are essential for analysis. The diseases are mostly on leaves and on stem of plant. The diseases are viral, bacterial, fungal, diseases due to insects, rust, nematodes etc. It is important task for farmers to find out these diseases as early as possible.

A. Objective

- Plant diseases cause a major production and economic losses in the agricultural industry.
- Image processing plays a crucial role in the detection of plant diseases since it provides best results and reduces the human efforts.
- The image processing could be used in the field of agriculture for several applications. It includes detection of diseased leaf, stem or fruit, to measure the affected area by disease, to determine the color of the affected area.
- These systems monitor the plant such as leaves and stem and any variation observed from its characteristic

features, variation will be automatically identified and also will be informed to the user.

B. Overview

Digital image processing is always an interesting field as it gives improved pictorial information for human interpretation and processing of image data for storage, transmission, and representation for machine perception. Image Processing is a technique to enhance raw images received from cameras/sensors placed on satellites, space probes and aircrafts or pictures taken in normal day-to-day life for various applications. This field of image processing significantly improved in recent times and extended to various fields of science and technology. The image processing mainly deals with image acquisition, Image enhancement, image segmentation, feature extraction, image classification etc.

II. LITERATURE SURVEY

A. Leaf diseases detection and prevention using image processing using matlab

Prajakta Mitka, et al [1] was proposed, "Leaf diseases detection and prevention using image processing using Matlab". In this paper, the farmers and agro help center use the different new technology to enhance the agriculture production. There are several diseases that affect plants with the potential to cause economic and social losses. Many of disease are most popular where disease spots occur on the sugar cane plant leaves. By this paper, we can find out particular disease using Digital image processing which helps to find disease and provide prevention for particular disease

In this paper, digital world research are continuously trying to increasing the collective of plants. They were achieved by using the higher breed seeds and plants. But one problem still exist which is a major concern of the cultivation of crop and that is crop diseases and the pesticides problem in turn the country suffers the lack of cultivation of plant. In this paper, fungi-caused diseases in sugarcane are the most predominant disease that appears as spots on the leaves.

B. Image Processing Techniques for Detection of Leaf Disease

Arti N. Rathod, et al [2] discussed, "Image Processing Techniques for Detection of Leaf Disease". In this paper, the images are captured by digital camera, mobile and processed using image growing, then the part of the leaf spot has been used for the classification purpose of the train and test. Image analysis can be applied for the following purpose to detect diseased leaf, stem, and fruit to identify affected area by disease.

They proposed method consists of two phases to identify the affected part of the disease. Initially Edge detection based Image segmentation is done, and finally image analysis and classification of diseases is performed using our proposed Homogeneous Pixel Counting Technique for Cotton Diseases Detection (HPCCDD) Algorithm. Their goal of this research work is to identify the disease affected part of cotton leaf spot by using the image analysis technique. Their work is to find out the computer systems which analyze the input images using the RGB pixel counting values and identify disease.

C. Plant Disease Detection in Image Processing Using MATLAB

Sandesh Raut, et al [3] was proposed, "Plant Disease Detection in Image Processing Using MATLAB". In this paper, the manual monitoring of disease do not give satisfactory result as naked eye observation is old method that requires more time for disease recognition. So in this paper, they introduced a modern technique to find out disease related to both leaf and fruit. To overcome disadvantages of traditional eye observing technique, we used digital image processing technique for fast and accurate disease detection of plant. In this paper, they developed k-means clustering algorithm with multi SVM algorithm in MATLAB software for disease identification and classification.

This paper introduces a MATLAB based system in which we focused on both leaf & fruit diseased area and used image processing technique for accurate detection and identification of plant diseases. The MATLAB image processing starts with capturing of digital high resolution images. They also deals, healthy and unhealthy images are captured and stored for experiment. Then images are applied for pre-processing for image enhancement. In this paper, images are converted from RGB Color Space to L*a*b* Color Space in which the L*a*b* space consists of a luminosity layer 'L*', chromaticity layer 'a*' and 'b*'. All of the color information is in the 'a*' and 'b*' layers and colors are classified using K-Means clustering in 'a*b*' space.

D. Image Processing Based Leaf Rot Disease, Detection of Betel Vine

Amar Kumar Dey was proposed [5], "Leaf rot disease detection for betel vine (Piper betel L.) Based on image processing algorithm". In this paper, the measurement of plant features was a fundamental element of plant science research and related applications. This information was related to plant features which was especially useful for its applications in plant growth modeling, agricultural research and on farm production. They have done the present investigation, that the image processing algorithms were developed to detect leaf rot disease by identifying the color feature of the rotted leaf area. Subsequently, the rotted area was segmented and area of rotted leaf portion was deduced from the observed plant feature data. Their results showed a promising performance of the automatic vision-based system in practice with easy validation. This paper describes the steps to achieve an efficient and inexpensive system

acceptable to the farmers and agricultural researchers as well for studying leaf rot disease in betel vine leaf. The highly sustainable crops are capable of generating high revenue and are of greater importance for national economy. Betel vine garden was subjected to high yield loss when it was attacked by Leaf rot. Nearly 30-100% leaf yield loss was reported due to leaf rot disease in betel vine. Thus, a sincere attempt was required to overcome the adverse effects of leaf rot on betel vine cultivation.

E. *Detection of plant leaf diseases using image egmentation and soft computing techniques*

Vijai Singh, et al [6] deals, "Detection of plant leaf diseases using image segmentation and soft computing techniques". In this paper, detection of plant disease through some automatic technique was beneficial as it reduces a large work of monitoring in big farms of crops, and at very early stage itself it detects the symptoms of diseases i.e. when they appear on plant leaves. This paper presents an algorithm for image segmentation technique which was used for automatic detection and classification of plant leaf diseases. They covers survey on different diseases classification techniques that can be used for plant leaf disease detection which was an important aspect for disease detection in plant leaf disease, was done by using genetic algorithm.

The existing method for plant disease detection was simply naked eye observation by experts through which identification and detection of plant diseases was done. For doing so, a large team of experts as well as continuous monitoring of plant was required, which costs very high when we do with large farms. At the same time, in some countries, farmers do not have proper facilities or even idea that they can contact to expert. Here, Automatic detection of the diseases by just seeing the symptoms on the plant leaves makes it easier as well as cheaper. This also supports machine vision to provide image based automatic process control, inspection, and robot guidance. Many different ways of performing image segmentation, ranging from the simple thresholding method to advanced color image segmentation methods were used. These parts normally correspond to something that humans can easily separate and view as individual objects. Computers have no means of intelligently recognizing objects, and so many different methods have been developed in order to segment images. The segmentation process is based on various features found in the image. This might be color information, boundaries or segment of an image. We use Genetic algorithm for color image segmentation.

F. *Leaf disease detection using image processing*

Sujatha R [7] was proposed, "Leaf disease detection using image processing" requires huge amount of work, knowledge in the plant diseases, and also require the more processing time. So we can use image processing for identification of leaf disease in MATLAB. Few steps also involved in identification of disease like loading the image, contrast enhancement, converting RGB to HSI, extracting of features and SVM.

They describes that the identification was done manually by the experienced people but due to the so many environmental changes the prediction was becoming tough. So we can use image processing techniques for identification of plant disease.

In this paper, generally we can observe the symptoms of disease on leafs, stems, flowers etc. so here they use leafs for identification of disease affected plants. The Study and analysis of cotton leaf disease detection using image processing work was carried on using k-means Clustering algorithm. The k-means concept which was added to the proposed system will divide the leaf into different clusters. The Comparison of different detection technique of leaf disease detection was mentioned. SVM and k-means clustering has been used by this system.

In Image Acquisition, we need to select the plant which was affected by the disease and then collect the leaf of the plant and take a snapshot of leaf and load the leaf image into the system. In Segmentation, the representation of the image was more meaningful and easy to analyze and the digital image was partitioned into multiple segments which can be defined as super-pixels. K-means clustering algorithm was used to cluster/divide the object based on the feature of the leaf in to k-number of groups. In Support Vector Machine they aims at gaining knowledge, making decisions from a set of data.

G. *Identifying multiple plant diseases using digital image processing*

Jayme et al [8] deals, "Multiple plant diseases identify the digital image processing". In this paper, human sight and cognition are remarkably powerful in identifying and interpreting patterns. Ambiguities may be resolved by laboratorial analysis, however this is a process that is often time consuming and expensive. This paper describes that, many producers around the world do not have access to technical advice from rural extension, making their crops especially vulnerable to yield losses and further problems caused by plant diseases.

From this paper, they have an algorithm which was divided into two branches, training and core. The core part uses the parameters and values determined in the training part to perform the disease identification. In other words, the first branch is used only when some training is necessary (for example if a new disease is to be included), while the second branch is the actual disease classifier, being the one with the potential to be used in practice.

III. METHODOLOGY

Diseases in plants cause major production and economic losses in agricultural trade worldwide. observation of health and detection of diseases in plants and trees is vital for property agriculture. To the most effective of our data, there's no device commercially accessible for period assessment of health conditions in trees. The classification strategies are often seen as extensions of the detection strategies, however rather than attempting to observe just one specific sickness amidst totally different conditions and symptoms, these ones attempt to determine and label whichever pathology has effects on the plant.

A. *Image Acquistion*

The images of the plant leaf are captured through the camera. This image is in RGB (Red, Green and Blue) form. Color transformation structure for the RGB leaf image is created, and then, a device-independent color space transformation for the color transformation structure is applied.

B. *Image Preprocessing*

To remove noise in image or other object removal, different pre-processing techniques is considered. Image clipping i.e. cropping of the leaf image to get the interested image region. Image smoothing is done using the smoothing filter. Image enhancement is carried out for increasing the contrast. The RGB images into the grey images using color conversion using equation.

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

(R-Red, G-Green, B-Blue)

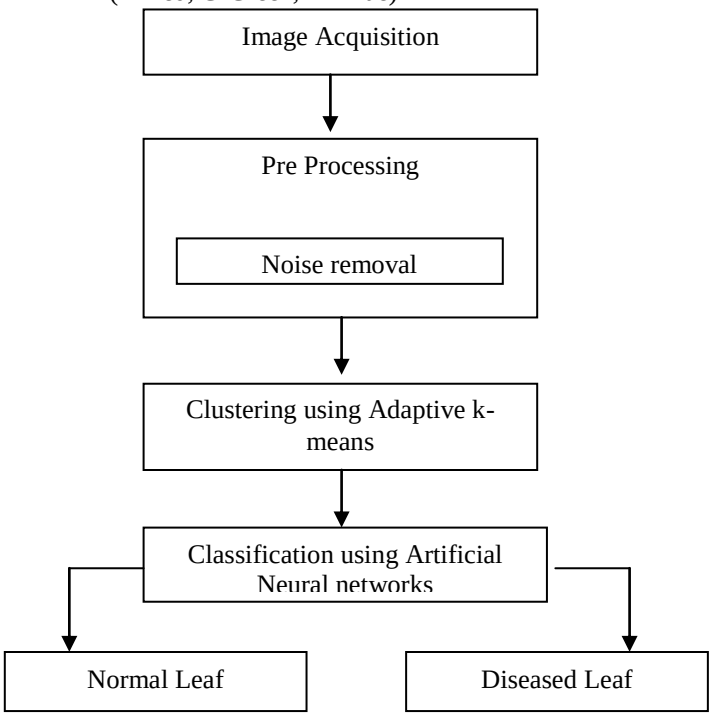

Fig. 1 Block Diagram for identifying and classification of plant diseases

C. *Image segmentation*

Segmentation means partitioning of image into various part of same features or having some similarity. The segmentation can be done using various methods like k-means clustering, converting RGB image into HIS model etc.

D. *Gradient Descent Algorithm*

The K-means clustering is used for classification of object 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.

1) The algorithm for K –means Clustering

1. Pick center of K cluster, either randomly or based on some heuristic.
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

E. *Feature Extraction*

Feature extraction plays an important role for identification of an object. In many application of image processing feature extraction is used. Color, texture, morphology, edges etc. are the features which can be used in plant disease detection.

1) Color co-occurrence Method

In this method both color and texture are taken into account to get an unique features for that image. For that the RGB image is converted into the HSI translation.

$$H=\begin{cases} \text{Theta} & \text{if } B<G \\ 360\text{-Theta}, & B>G \end{cases} \quad \text{- - - - - (2)}$$

$$S=1-3/(R+G+B)\ [\min(R,G,B)] \quad \text{- - - - - - (3)}$$

$$I=1/3\ (R+G+B) \quad \text{- - - - - (4)}$$

(H-Hue, S-Saturation, I- Intensity)

IV. RESULTS

A. *Dataset*

The k-means algorithm is tested on two different data sets. One is the Cercospora leaf and the other is healthy leaf, below and the other are taken from the database, such that the name of the disease will be either bacteria, virus and fungal and the accuracy ranges from 90 to 96.

Figure 1:Cercospora

Figure 2: Healthy leaf

B. Experiment & Results

In our Experiment, for input data disease, samples of leaves like betel leaf with bacterial disease, are considered. The original images which are followed by output segmented images. Segmented image can be classified into different leaf diseases. The input and output image where input image is a betel leaf with early Cercospora disease and output image shows the classification of disease using feature extraction method. In the same manner classification of diseases of other input plant leafs .The method is first done using K-Mean Clustering and shows its efficiency with accuracy

The numbers of leaf disease samples that were classified into three classes of leaf disease using proposed k –mean algorithm. From the results it can be seen that only few samples from betel leaf spot and bacterial leaf spot leaves.

	1	2	3	4	5	6
1	0.0789	0.9783	0.7626	0.9749	14.8439	47.8117
2	0.4668	0.8657	0.7967	0.9592	14.1501	48.1396
3	0.3676	0.9102	0.7573	0.9625	16.4441	51.4194
4	0.5412	0.7510	0.5382	0.9222	17.9717	37.6635
5	0.5128	0.7103	0.8947	0.9717	17.1185	35.5205
6	0.6976	0.8739	0.4873	0.9104	31.5604	56.4596
7	0.4886	0.9580	0.2687	0.9403	71.8528	83.0729
8	0.4309	0.8966	0.7660	0.9656	17.4376	52.4639
9	0.5761	0.9092	0.7104	0.9584	23.8136	60.2088
10	0.7462	0.9098	0.5279	0.9007	40.0473	73.8575
11	0.8894	0.8263	0.8185	0.9651	16.4181	55.6534
12	0.4140	0.9702	0.4106	0.9730	76.6184	97.9821

Figure 3: Training dataset

C. Summary of Results

When the segmentation method integrated with K-means clustering algorithm and linear discriminant analysis was used, the segmentation effects for the sub-images of the four leaf diseases were best. The segmentation results of the sub-images of the three betel leaf diseases using the segmentation method integrated with K-means clustering algorithm and linear discriminant analysis. Using this segmentation method, all lessons in the original sub-images were effectively segmented. The results indicated that this segmentation method could effectively implement the automatic segmentation of sub images of the three betel leaf diseases. Therefore, the image segmentation was implemented using the segmentation method integrated with K-means clustering algorithm and linear discriminant analysis for further feature extraction, feature normalization, feature selection and building of disease recognition models.

V. CONCLUSION

In this project we have proposed feature extraction based concept of detecting disease of Betel leaf. After doing review on various techniques and algorithms we had come to conclusion that, k-means algorithm gives the better result as compare to other algorithms. This approach can also be developed using normal techniques like JAVA, but using Matlab gives the efficient and effective result. As the main focus of this application is user-friendly, this application id designed in such a way that it supports Multi-Lingual concept. This application is helpful for farmer and laboratory where they are can easily protect their crops and there will be increase in growth of production.

REFERENCES

[1] P. P. Prajakta Mitkal, "Leaf Disease Detection and Prevention Using Image Processing using Matlab," International Journal of Recent Trends in Engineering & Research, vol. 2, no. 2, p. 26, 2016.

[2] B. T. Arti N. Rathod, "Image Processing Techniques for Detection of Leaf Disease," International Journal of Advanced Research in Computer Science and Software Engineering, vol. 3, no. 11, p. 397, 2013.

[3] A. F. Sandesh Raut, "Plant Disease Detection in Image Processing Using MATLAB," International Journal of Innovative Research in Science, Engineering and Technology , vol. 6, no. 6, p. 10373, 2017.

[4] M. M. K. Rajneet Kaur, "A Brief Review on Plant Disease Detection using in Image Processing," International Journal of Computer Science and Mobile Computing , vol. 6, no. 2, p. 101, 2017.

[5] M. S. Amar Kumar Deya, "Image Processing Based Leaf Rot Disease, Detection of Betel Vine," International Conference on Computational Modeling and Security, vol. 1, no. 1, p. 1, 2016.

[6] A. M. Vijai Singh, "Detection of plant leaf diseases using image," INFORMATION PROCESSING IN AGRICULTURE, vol. 4, no. 1, p. 41, 2017.

[7] Y. S. K. Sujatha R, "Leaf disease detection using image processing," Journal of Chemical and Pharmaceutical Sciences, vol. 10, no. 1, p. 670, 2017.

[8] J. G. A. Barbedo, "Identifying multiple plant diseases using digital," Elsevier, vol. 1, no. 1, p. 104, 2016.

[9] Shiddalingappa Kadakol, "Intelligence System for Leaf Extraction and Disease Diagnostic," International Journal of Computer Sciences International Journal of Computer Sciences &Engineering, vol. 4, no.3 , p. 2016.

[10] **Asha Patil**, " Leaf Disease detection using Image Processing Techniques," International Journal of Scientific Research in Computer Science and Engineering, vol. 6, no. 1, p. 33-36.

