Plant Leaf Disease Segmentation and Feature Extraction using Image Processing

Khushal Khairnar 1, Sajjidullah Khan 2

1 (Computer Science and Engineering, Research Scholar, Sandip University, Faculty, COEP, India)

2 Associate Professor(School of Computing Science & Engineering, Sandip University, Nashik, India)

ABSTRACT

Plant leaf diseases are decreasing growth and production of plants. At present, farmers are identifying, diagnosing diseases and monitoring health in plants by their own knowledge and experience. Naked eye observation by farmers and experts on big plantation areas cannot be possible each time and it can be expensive. Accurate identification of visually observed diseases, symptoms and controls has not studied yet. Therefore a fast automatic, economical and accurate system is an essential research topic that may improve in leaf disease detection of plant disease. The proposed automatic early leaf disease segmentation and feature extraction system is based on image processing and machine learning where segmenting the major diseases such as Bacteri

al Blight, Alternaria leaf spot and Cercospora leaf spot. Initially, the infected leaf images are captured from cotton plant fields by using a digital camera. Scaling, background removing and color conversion are done in the preprocessing phase. After preprocessing, the infected region is obtained by using K-means algorithm and Thresholding technique. Color, texture and edge features are extracted by using image processing technique.

Keywords : *clustering, feature extraction, image processing, segmentation.*

I. INTRODUCTION

Cotton is the most important crop for Indian economy. Cotton has most globally significance in oil and protein yielding as well as making fiber. India is the second largest producer of cotton in the world and produced 6.71 million metric tons of cotton. Automatic early disease detection of cotton plant may be useful for monitoring large fields of cotton crops and also prevent production losses. There are more than 80% diseases of cotton plant on its leaves. So it is easy to capture images of cotton leaves [10]. The Cotton plant leaf has diseases like bacterial, fungal,viral and due to insects. The major diseases on cotton plant leaf are leaf spot, Alternaria, Bacterial blight and leaf curl. Computer is an important device in the agriculture application for image processing. First healthy and infected images are captured through the camera. Image preprocessing is required for removing noise and smoothing of the images. Image segmentation is an important technique for segmenting the images into smaller segments that are more meaningful. Images are segmented based on the similar properties of objects like texture, shape and color. In this system, infected region of the leaf diseases are segmented and different features such as color, texture and shape are calculated.

II. PROPOSED SYSTEM

The proposed system for leaf disease segmentation and feature extractions using image processing consist of following steps:

1. Cotton plant leaf image acquisition
2. Image preprocessing of acquired images
3. Color based Segmentation
4. Edge Detection
5. Feature Extractions

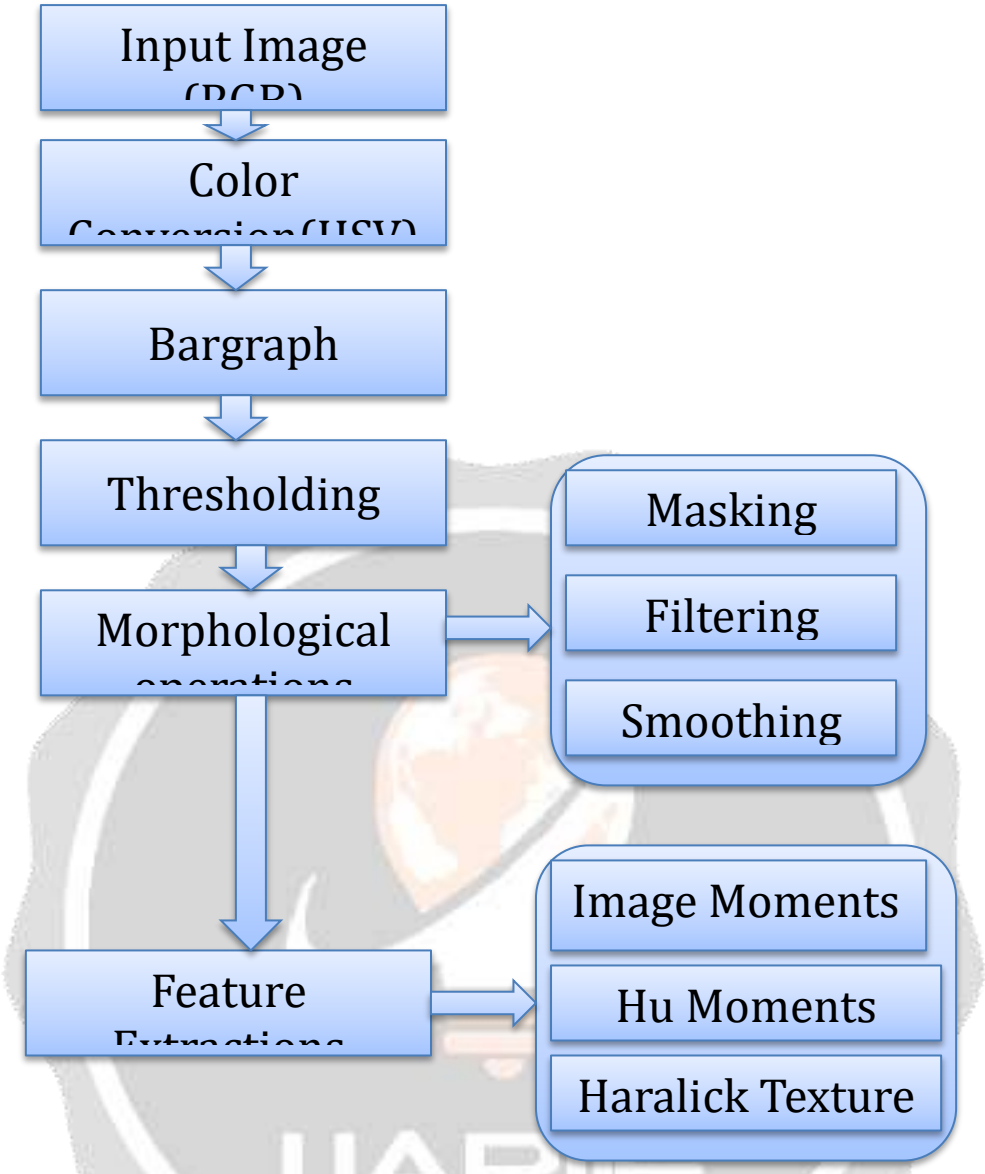

Fig 1: the flow diagram of proposed system

III.I Cotton plant leaf image acquisition:
Initially, the infected leaf images are captured through digital camera with 300*500 resolutions from the region of Maharashtra, India. By default all images are saved in the RGB color space model while capturing the images from a digital camera. Some of the images were downloaded from the PlantVIllage database. The size of the dataset is 5000 leaf images of four classes including healthy images[10]. The dataset is summarized per classes as followed:

III.I.I Bacterial Blight (Angular leaf spot): It is bacterial disease, where more than 10% losses on cotton plant yielding are due to bacterial blight. It is affected by Xanthomonas citri bacteria. The red or brown border is shown on lesion of leaf. It is also infected on stem and petiole of cotton plant. Angular shape is progressed on lesions of upper leaf surface.

Fig. 2. Bacterial Blight on upper part of leaf

Fig. 3. Bacterial Blight on lesion of leaf

III.I.II Alternaria Leaf Spot: It is one the major disease which is observed on all cotton plantation countries in the world. This foliar disease is highly infected on cotton production and reduced quality and quantity. Brownish spot is shown on lower leaf surface and on boll of plant.

Fig. 4. Alternaria leaf spot

III.I.III Cerespora leaf spot: This is also foliar disease caused by Cercospora gossypina. Raddish spot is shown on middle of leaf. It is usually affected with stress such as drought and nutrient deficiency.

Fig. 5. Cerespora leaf spot

III.II Image preprocessing of acquired images:

Image preprocessing operations such as resizing with 100*100 resolution, color transformation, filtering, removing noise and background are required for finding meaningful contents from infected images. To find the infected region, It is necessary to convert RGB color space into the device-independent color space. In HSV color space, H is hue the color portion of the model. S is the saturation, describes the amount of grey in a particular color, V indicates value, works in a conjunction with saturation and describes the brightness or intensity of the color [1]. Python (3.7) is used to implement all proposed algorithms in this work. The PyCharm Community-IDE 2020 is used for implementation.

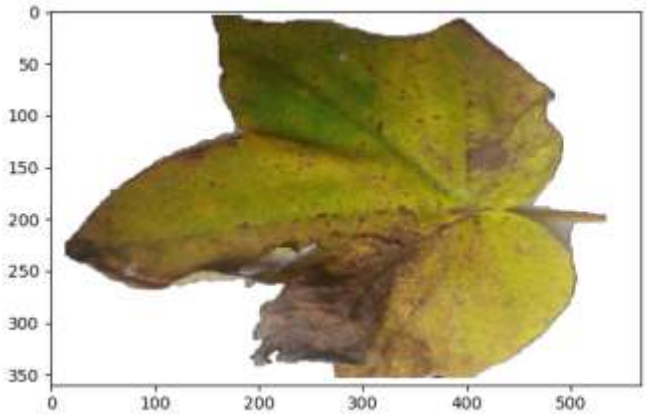

Fig. 6. Original RGB Image

After RGB to HSV color conversion shown in fig 7, each HSV image is observed that hue describes the clear perception of cotton infected leaf area also Hue is useful for masking the background and rest of the infected leaf image

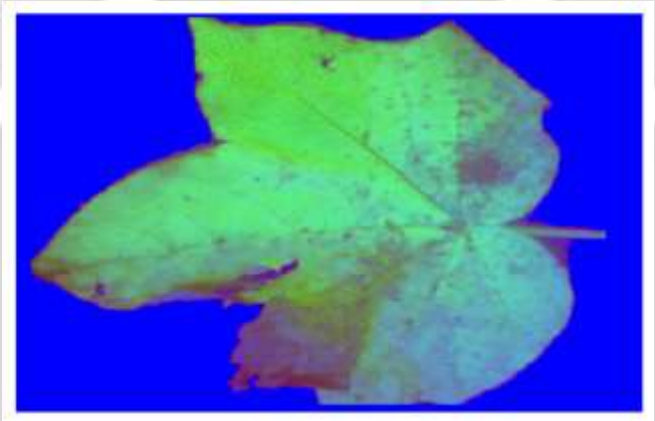

Fig. 7. HSV Image

III.III Color based Segmentation and Smoothing of the image:

In the Image segmentation phase, it is the process of segmenting an image into multiple segments, often based on characteristics of the pixels in the image. Here, K-mean algorithm is used for segmentation. Color based segmentation is used to find infected regions that describe homogeneous colors in the infected leaf image correspond to separate clusters. Cluster is a class of pixels based on similarities in color is shown in fig 9. Smoothing is used to reduce noise within HSV image. Smoothing is performed by using MeanShift Filtering method to enhance and improve quality of image as shown in fig 8. K-Means clustering apply on smooth image to find infected region, it generates four color based cluster(k=4) as shown in bargraph. K –Mean clustering is used and based on the Euclidean distance formula (minimizing the sum of squares of distances between the objects and the corresponding centered classes) [2]

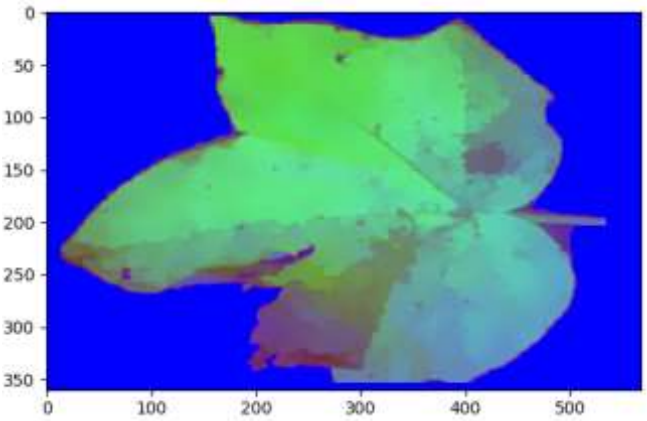

Fig. 8. Segmented Image

Fig. 9. Bargraph

III.IV Edge Detection

Edge Detection is an important technique for finding the boundaries of objects within the infected images. Canny Edge detection algorithm is used for finding the boundaries in the form of tooth, smooth and wavy. It also finds the midrib alignment and vein pattern of infected leaf images.

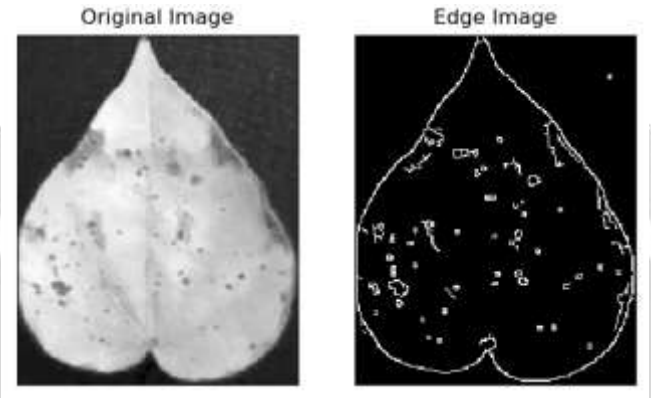

Fig. 10. Edge Detection

III.V Feature Extraction

After segmentation, set of features are extracted to describe the infected region. In this step, significant features are extracted such as color, texture and shape and these features can be used to determine the meaning of a given image.

III.V.I Image moments:

Image moments are a weighted average of image pixel intensities. The simplest kind of moment we can define is given below:

$$M = \sum_{x} \sum_{y} I(x, y) \qquad (1)$$

The above equation is calculating the sum of all pixel intensities. All pixel intensities are weighted only based on their intensity.

III.V.II Hu Moments:

Calculating moments that are invariant to translation, scale, and rotation are called Hu Moments. These are a set of 7 numbers calculated using central moments that are invariant to image transformations. The first 6 moments have been proved to be invariant to translation, scale, and rotation, and reflection. While the 7th moment's sign changes for image reflection.

III.V.III Haralick texture:

Haralick texture features are calculated from a Gray Level Co-occurrence Matrix, (GLCM), a matrix that counts the co-occurrence of neighboring gray levels in the image. The GLCM is a square matrix that has the dimension of the number of gray levels N in the region of interest (ROI). Haralick described 14 statistics that can be calculated from the co-occurrence matrix with the intent of describing the texture of the image. 532 statistical features are calculated for each disease and healthy image. All these features are stored in H5PY file for further process such as classification and detection of disease.

III. CONCLUSION

Farmers judge the diseases by their own experience and knowledge; it results in misidentification of disease. So an automatic leaf disease segmentation and feature extraction is an expert system which is helpful for finding the infected region and its statistical features. The major diseases like Bacterial Blight, Cercospora Leaf Spot, Alternaria Leaf Spot are automatically classified based on image preprocessing technique. This system also applied for other plant diseases detection.

REFERENCES

[1] Khushal Khairnar, Nitin Goje, "Image Processing Based Approach for Diseases Detection and Diagnosis on Cotton Plant Leaf", Techno-Societal2018, 2020, Springer.
[2] Siddharth Singh Chouhan, Ajay Kaul, "Bacterial Foraging Optimization Based Radial Basis Function Neural Network (BRBFNN) for Identification and Classification of Plant Leaf Diseases: An Automatic Approach Towards Plant Pathology", Digital Object Identifier 10.1109/ACCESS.2018.2800685, IEEE ACCESS 2018.
[3] Hilman F. Pardede, EndangSuryawati, Rika Sustika, and Vicky Zilvan"Unsupervised Convolutional Autoencoder-Based Feature Learning for Automatic Detection of Plant Diseases", International Conference on Computer, Control, Informatics and its Applications, 158-162,IEEE-2018.
[4] Vijay Kumar V, Vani K S "Agricultural Robot: Leaf Disease Detection and Monitoring the Field Condition Using Machine Learning and Image Processing", ISSN 0973-1873 Volume 14, Number 7 (2018), pp. 551-561,International Journal of Computational Intelligence Research-2018.
[5] Wan-jie Liang, Hong Zhang, Gu-feng Zhang & Hong-xin Cao "Rice Blast Disease Recognition Using a Deep Convolutional Neural Network", Scientific Report-2018.
[6] Melike Sardogan and YunusOzen, "Plant Leaf Disease Detection and Classification Based on CNN with LVQ Algorithm", 3rd International Conference on Computer Science and Engineering, IEEE-conference, 2018.
[7] Sharath D M, Akhilesh, S Arun Kumar, Rohan M G and Prathap C, "Image based Plant Disease Detection in Pomegranate Plant for Bacterial Blight", International Conference on Communication and Signal Processing, April 4-6, 2019, India.
[8] Fazeel Ahme'd Khan, Adamu Abubakar Ibrahim and Akram MZeki "Environmental monitoring and disease detection of plants in smart green house using internet of things", IOP Publishing, Journal of Physics 2020.
[9] http://cotcorp.gov.in/current-cotton.aspx?pageid=4.