Image based Identification of Leaf Crumple and Leaf Spot Diseases in Cotton Plant

Anush Reddy Kommareddy, Satya Anirudh Polisetty, Chetan Sai Kurra, S. Padmavathi, Mithun Chandra Pokuri

Abstract--- Identification of plant diseases based on images derived from computer vision is a major requirement for smart agriculture. Conventional algorithms warrant large dataset for better accuracy. They perform well with large variation in color or explicit probes on a specific disease. This paper considers 4 major diseases of cotton plants with a combination of images with and without color variation. This paper adapted image processing algorithms to extract precise features for classification, highly preferred and apt, when the dataset sizes are limited. Verification results of the proposed method validate its rationale and viability.

Keywords--- Cotton Plant Disease Identification, Leaf Disease Feature Extraction, Color Image Enhancement, Histogram Equalization, Color Edge Detection.

I. INTRODUCTION

India is an agricultural country, 70% of whose economy depends on agriculture. Both the quality and quantity of agricultural products are significantly affected by plant diseases significant reduction in. Diseases chiefly ensure from unassailable weather patterns such as excessive and or untimely rainfall, drought, and drastic changes in temperature as well as in apt maintenance. Insect infestations and pernicious side-effects of pesticides also contribute to this nasty problem. Getting a firm handle on plant diseases, manifested by unsettling loss of major crops, is an irrevocable factor in agro-economies of countries like India. Cotton is one of the most important fiber crops in India that provides basic raw materials for the textile industry. There are mainly four cultivated species of cotton: Gossypium Arboreum, G.herbaceum, G.hirsutum, and G.barbadense. Timely detection of cotton plant diseases at an earlier stage is inevitable to control the loss of major crops. Widespread technological advances, in the prevailing era of the digital age, afford the adoption of computer vision for the identification of plant diseases. Plant researchers affirm that about 80 % of the cotton plant diseases can be traced to their leaves. Therefore, this paper is focused on probe of the leaf rather than the whole cotton plant. Besides, the leaves, conspicuous by the cotton plantations, are accessible for visual examination and analysis of any disease-symptoms, obviating damage to the plant.

Diseases emanate from the invasion of leaf tissues by disease-causing agents such as bacteria, virus, and fungus, leading to degradation of the leaf as well as the plant. Undesirable symptoms can be characterized by random spots on the leaves, changes in the color, and shape of cotton-plant leaves. There 4diseases are primarily related to cotton leaves i.e. Myrothecium Leaf Spot, Alternaria leaf Spot, Leaf Crumple, Cercospora leaf spot. Automated classification of cotton leaf diseases, using the computer-vision technology, involves the following steps: Acquisition of leaf images, pre-processing, segmentation, feature extraction, and classification. Resume In this paper, images that contain leaf partition of cotton plant was pre-processed for removal of noise, the leaf regions were segmented using Otsu, and features were extracted from the segments. The derived features were then classified byK-Nearest Neighbours (KNN), Naive Bayes and Decision tree algorithms for. Results of these classifiers were analyzed and summarized. The rest of the paper is organized as follows. Literature survey is discussed in Section II. Section III, explains the methodology adopted in this study. The results are presented in Section IV, followed by concluding remarks in Section V.

II. LITERATURE SURVEY

The basic steps entailed in the detection of leaf disease, using image processing, include the following: image acquisition, image pre-processing, image segmentation, feature extraction, detection and classification of plant disease.

Patki [1] segmented the regions of leaves with disease, from an Red Green Blue (RGB) image by Otsu's global thresholding method. Shape and color features were extracted, and classified with the Support Vector Machine (SVM) classifier. Sivasangari [2] developed an Android application, through which leaf images were captured and converted to Hue Saturation Intensity (HSI) image, through color transformation. Then, edge detection operators such as erosion, dilation were applied, followed by SVM classification, with a genetic algorithm.

Manuscript received June 10, 2019.

Anush Reddy Kommareddy, Dept of Computer Science and Engineering, Amrita School of Engineering, Coimbatore, T.N, Amrita Vishwa Vidyapeetham, India. (e-mail: cb.en.u4cse15022@cb.students.amrita.edu)

Satya Anirudh Polisetty, Dept of Computer Science and Engineering, Amrita School of Engineering, Coimbatore., T.N, Amrita Vishwa Vidyapeetham. India. (e-mail: cb.en.u4cse15035@cb.students.amrita.edu)

Chetan Sai Kurra, Dept of Computer Science and Engineering, Amrita School of Engineering, Coimbatore, T.N, Amrita Vishwa Vidyapeetham, India. (e-mail: cb.en.u4cse15013@cb.students.amrita.edu)

S. Padmavathi, Dept of Computer science and Engineering, Amrita School of Engineering, Coimbatore, T.N, Amrita Vishwa Vidyapeetham, India. (e-mail: s_padmavathi@cb.amrita.edu)

Mithun Chandra Pokuri, Dept of Computer science and Engineering, Amrita School of Engineering, Coimbatore, T.N, Amrita Vishwa Vidyapeetham, India. (e-mail: cb.en.u4cse15034@cb.students.amrita.edu)

Color, edge, and shape were the features used for the analysis. The paper claims a recognition accuracy of 93 %. Chaudhari [3] suggested k-mean clustering with Discrete Wavelet Transform (DWT) for leaf image segmentation and classification, using a neural network. (NN)[4] adopted the deep learning algorithm for classification; the paper claimed that accuracy was improved by increasing the layers of the NN.SVM classifier [5] was engaged for classification between normal leaf and diseased leaf; the algorithm was limited to 3diseases. [7], suggested that a user can realize the affected area of a leaf in terms of percentage, if the disease is correctly identified. In [8] after detection of the disease, the name of pesticide will be sent to our mobile with the help of GSM. Krithika [11] adapted multi-class SVM for classification of cucumber leaves. Srunitha [12] classified mango leaves with artificial NN and SVM. [9] extracted and classified features using Histogram Equalization, in an artificial NN, and back propagation in a NN. [10] used NN to identify plant diseases. Premalatha[6] used the median filter to reduce the noise in images, and the spatial fuzzy c-means clustering algorithm for segmentation. The images were classified by the spatial probabilistic neural network.

III. METHODOLOGY

The leaf spots are the key indicators for the above mentioned group of plant diseases. Since leaf spots can be identified from their colour, simple segmentation algorithms is adequate to verify their identity. The categories of Cercospora, Myrothecium, and the Alternaria leaf spots can be identified by the number of leaf spots and their sizes.

Leaf spot segmentation and feature extraction

Images of Cotton leavesfor different diseases, at various resolutions, are stored in the dataset. An image of a cotton leaf is loaded from the dataset. During pre-processing, the contrast of the image is enhanced. The RGB image of cotton leaf is converted to LAB color image, which is subsequently converted into Black and White (BW) image. Morphological closing is applied to the complemented BW image using the structural element as specified by (1)

$$A \bullet B = (A \oplus B) \Theta B \quad (1)$$

Where,

A is cotton disease leaf and B is disk structural element. Closing of an image A is dilation of A by B, which is followed by erosion on that image, by B.

The features from a given cotton leaf image are extracted using the Connected components algorithm. The properties of eccentricity, major axis length, area of each connected component, are determined and stored as features. These features are then fed as inputs into the classifiers, for identification of the disease.

Classification

KNN classifier, Naive Bayes and Decision tree algorithms are used for classification. Given a set of sample leaf images, with a disease label, whose features are extracted and stored in the knowledge base, the input image is pre-processed, features are extracted and fed as input to the classifiers.

KNN classifier: The KNN Classifier computes the distance of the test sample with the knowledge base to determine k closest samples, from which the disease label is derived. In this classifier, Euclidean distance is used to find the distance between the features, P and Q. Let P and Q be denoted by feature vectors $P = (x_1, x_2, \ldots, x_m)$ and $Q = (y_1, y_2, \ldots, y_m)$, where m is the dimensionality of the feature space. The distance between P and Q is calculated by (2).

$$\text{distance}(P,Q)=\sqrt{\frac{\sum_{i=1}^{m}(x_i-y_i)^2}{m}} \quad (2)$$

a) *Bayes classifier:*The probability of a feature vector, belonging to a known sample, is calculated as posterior probability, and class labels are assigned accordingly.Posterior probability can be calculated by (3).

$$P(C_i|X) = \frac{P(X|C_i)}{P(X)} \quad (3)$$

where

X is the feature.

C_i is the class i.e. cotton leaf disease.

$P(C_i|X)$ is the posterior probability of class $C_i$conditioned on X.

$P(X|C_i)$ is the probability of X conditioned on class C_i.

P(X) is the prior probability of X.

Class label is assigned based on (4).

$$c_0 = \arg\max_{c_i} P(C_i|X) \quad (4)$$

b) *Decision tree classifier:*This is a hierarchical classification algorithm, in which the ranking of features using their information gain.This classifier assigns the class labelusing tree representation. Each internal node of the tree represents an attribute, and each leaf node represents a class label. Information gain is used as a measure to select an attribute . Information gain is calculated by(5).

$$\text{Info}(D) = -\sum_{i=1}^{m} p_i \log_2(p_i) \quad (5)$$

D is the data. p_i is the probability that a tuple in D belongs to class C_i.

Leaf crumple disease Identification

Since a leaf crumple does not have clear colour distinction from the normal region, a suitable algorithm has to be used to differentiate them, prior to segmentation and feature extraction. Histogram equalization on the colour component is performed to accentuate the distinction between a crumpled leaf and the parts of a normal leaf.

An RGB image of crumpled, diseased cotton leaf is converted into HSI image. Next, Histogram equalization is applied on the H and S parts of the HSI image.

Histogram equalization algorithm

1. Derive intensity mapping for image pixel data by finding the frequency of the intensity values.
2. Calculate the probability of occurrence of a pixel for all the gray scale levels by (6).

$$p_x(i) = \frac{n_i}{n} \quad 0 \le i < L \quad (6)$$

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

p_x(i) is the probability of occurrence of a pixel at level i.

n_i is the frequency in the level i.

n is the total frequency.

3. Calculate the cumulative distribution function for all the gray scale levels by (7).

$$cdf_x(i)=\sum_{j=0}^{i} p_x(j) \quad (7)$$

$cdf_x(i)$ is cumulative distribution function value at level i.

$p_x(j)$ is the probability of occurrence of a pixel at level j.

4. Then cdf is normalized to [0,255] i.e. histogram equalization is done by (8).

$$h(k)=\frac{cdf(k)-cdf_{min}}{(mxn)-cdf_{min}}*(L-1) \quad (8)$$

$cdf(k)$ is cumulative distribution function value at level k.

cdf_{min} is minimum cumulative distribution function value.

(mxn) is number of pixels in image.

L is total number of gray levels.

Otsu segmentation is implemented on the histogram equalized image by computing histogram and probabilities for each intensity level. Then edge detection algorithm is applied on the segmented image.

IV. RESULTS

The dataset which was acquired for this experimentation contains 25 images for each of the 4 diseases namely Alternaria leaf spot, Cercospora leaf spot, Leaf crumple and Myrothecium leaf spot. To show the importance of the proposed methodology, the images are classified, based on the pixel values, using different classifiers. Images in the dataset were resized to the fixed size (64x64 pixels). In this dataset, 70% of the data was used as a training set, and the residual 30% was used as a testing set. Three classifiers-KNN, Naive Bayes and Decision tree were applied to the dataset. Accuracy of these classifiers is shown in Table 1.

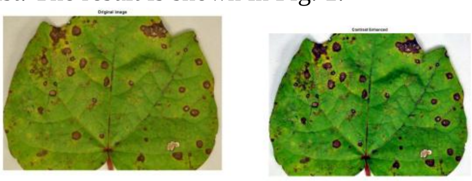
TABLE I. Classification Output

Classifiers	Accuracy
KNN	35.29%
Naive Bayes	21.43%
Decision tree	36.36%

Observation of the results make plain that all of the classifiers performed poorly with accuracies of less than 50%, when pixels are used directly. So, in order to improve the accuracy of classification, segmentation and feature extraction are to be applied.

Results of cercospora leaf spot disease

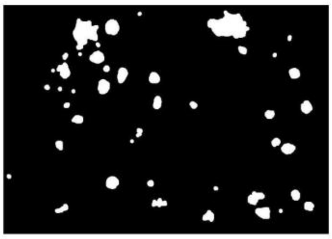
Cotton leaf with Cercospora leaf spot disease was considered as input and pre-processed by enhancing contrast. The result is shown in Fig. 1.



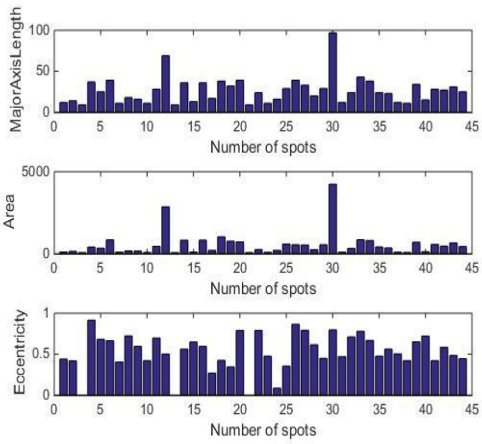
(a) (b)

Fig. 1.(a)Original Image (b)Pre-processed image.

RGB image in Fig.1. (b) was converted to LAB colour image. Components that were attached to the border in binary image were removed, using morphological operations, as shown in Fig.2. (a).The number of spots was calculated and features were extracted using Connected components algorithm. A graph was plotted between the number of spots and the features of the components (Eccentricity, Major axis length and Area),as in Fig.2. (b). Similar processing was conducted for other diseases and the results are shown in Fig. 3 and Fig. 4.



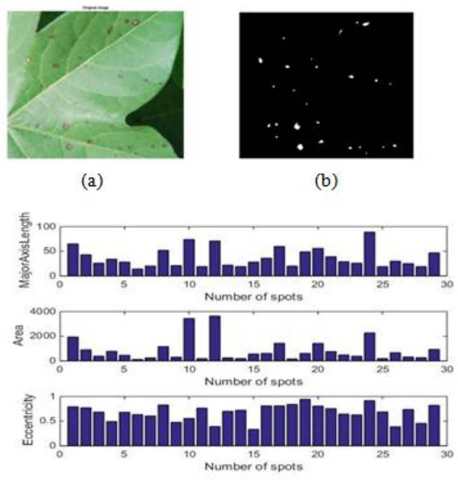
(a)



(b)

Fig. 2.(a)Output Image (b)Plot between number of spots and features.

Results of Alternaria leaf spot disease



(c)

Fig. 3.(a)Original image (b)Output image (c)Plots between number of spots and features

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

Results of Myrothecium leaf spot disease

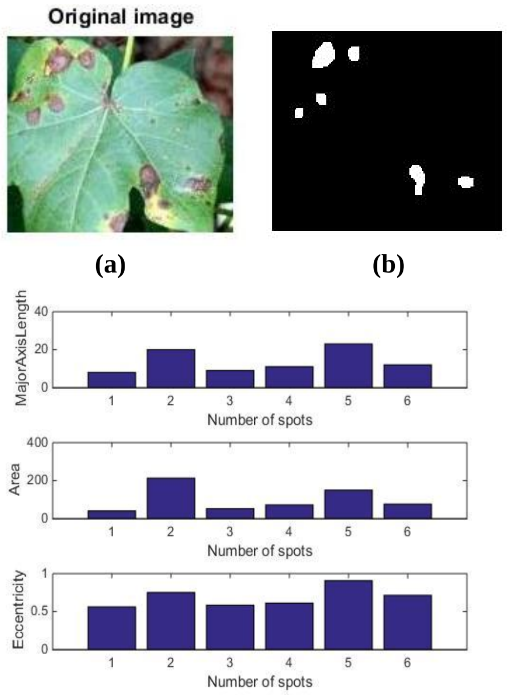


Fig. 4.(a)original image (b)Output image (c)Plots between number of spots and features

Results for leaf crumple disease

Cotton leaf image with leaf crumple disease is taken as input and converted to HSI image. Otsu segmentation was applied to the combined image of H and S followed by edge detection on the segmented image.

The result is shown in Fig. 5.

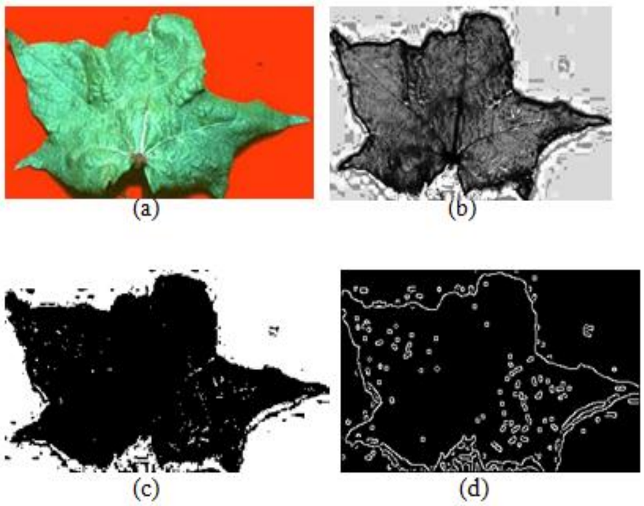


Fig. 5.(a)Original Image(b)H&S part of HSI image(c)segmented image (d)Applying edge detection on segmented image

From the features extracted it could be identified that the classification of the leaf diseases could be done with better accuracy.

V. CONCLUSION

Smart agriculture can very well be accomplished in conjunction with the steady growth of computing and related technologies. Identification of the plant disease, using computer vision and image processing, plays a crucial role in the digital age. Neural network based algorithms demand large datasets for training, which may not be available in many cases. Standard machine learning algorithms require proper selection of features for achieving better accuracy.

Disease identification without damaging the plant requires their recognition from the major plant part that is naturally exposed. In this paper a plant’s leaf was considered for disease identification. Disease without much colour variation in leaf is a challenging case for image processing algorithms. This paper tackled this challenge by using image enhancement techniques in an appropriate colour model. Simple features extracted from the leaves can be used for better performance of the classifiers.

REFERENCES

1. Shweta R. Astonkar1, Dr. V. K. Shandilya2 ,”Detection and Analysis of Plant Diseases Using Image Processing Technique ”, International Research Journal of Engineering and Technology, Volume: 05 Issue: 04 Apr-2018.
2. Ms.Nidhi Rajesh Savaji, Vrushabh khandelwal, Kalyani Bhawar, Punam Wankhede, Ravi Kiran Rajbhure,”Plant Disease Detection Using Image Processing”,IJCESR, Volume-5, Issue-3, 2018.
3. Rakesh Chaware, Rohit Karpe, Prithvi Pakhale, Prof.SmitaDesai4,”Detection and Recognition of Leaf Disease Using Image Processing ”,IJESC, Volume 7 Issue No.5 ,May-2017.
4. Pallavi.S.Marathe,” Plant Disease Detection using Digital Image Processing and GSM ”,IJESC, Volume 7 Issue No.4 April 2017.
5. Sujatha R, Y Sravan Kumar Garine Uma Akhil .” Leaf disease detection using image processing ”, Journal of Chemical and Pharmaceutical Sciences, Volume 10 Issue 1 January - March 2017.
6. A.Sivasangari,K.priya,K. Indira,” Cotton Leaf Disease Detection and Recovery Using Genetic Algorithm”, International Journal of Pure and Applied Mathematics, Volume 117 No.22, 2017.
7. Mr.Chandrakant Deelip Kokane ,Prof.N.L.Bhale,“To Detect and Identify Cotton leaf disease based on pattern recognition technique”,IJARIIE,Volume-3 Issue-4 ,2017.
8. Supriya S. Patki, Dr. G. S. Sable,”Cotton Leaf Disease Detection & Classification using Multi SVM ”, International Journal of Advanced Research in Computer and Communication Engineering, Vol. 5, Issue 10, October 2016.
9. Vijay S.Bhong , Prof.B.V.Pawar,”Study and Analysis of Cotton Leaf Disease Detection Using Image Processing”,IJARSET,Vol. 3, Issue 2 , February 2016.
10. Premalatha.V, Valarmathy.S, Sumithra.M.G,“Disease Identification in Cotton Plants UsingSpatial FCM & PNN Classifier”, International Journal of Innovative Research in Computer and Communication Engineering, Vol. 3, Issue 4, April 2015.
11. P.Krithika,S.Veni ,"Leaf Disease Detection on Cucumber Leaves Using Multiclass Support Vector Machine",2017.
12. K.Srunitha,D.Bharathi,"Mango Leaf Unhealthy Region Detection and Classification",2018.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication