Machine Learning Regression Technique for Cotton Leaf Disease Detection and Controlling using IoT

Adhao Asmita Sarangdhar[1] , Prof. Dr. V. R. Pawar [2]
Dept. of Electronics and Telecommunication
Bharati Vidyapeeth's college of Engineering for Women, BVCOEW
Pune, India
adhao.asmita87@gmail.com[1] , pawar.vijaya40@gmail.com[2]

Abstract— **Cotton is one of the most important cash crops in India. Every year the production of cotton is reducing due to the attack of the disease. Plant diseases are generally caused by pest insect and pathogens and decrease the productivity to large-scale if not controlled within time. This paper presents a system for detection and controlling of diseases on cotton leaf along with soil quality monitoring. The work proposes a Support Vector Machine based regression system for identification and classification of five cotton leaf diseases i.e. Bacterial Blight, Alternaria, Gray Mildew, Cereospra, and Fusarium wilt. After disease detection, the name of a disease with its remedies will be provided to the farmers using android app. The Android App is also used to display the soil parameters values such as humidity, moisture and temperature along with the water level in a tank. By using Android app farmers can ON/OFF the relay to control the motor and sprinkler assembly according to need. All this leaf disease detection system and sensors for soil quality monitoring are interfaced using Raspberry Pi which make it independent and cost effective system. The overall classification accuracy of this proposed system is 83.26%.**

Keywords—Cotton Leaf Disease; Color Transform; Gabor Filter; Median Filter; Color Moment; Support Vector Machine; Raspberry Pi; Android App; Sensors.

I. Introduction

Cotton is one of the most important cash crops in India and affects India's economy in many ways [4]. Large number of the population depends on Cotton crop either for its cultivation or for the purpose of processing [4]. It is observed that the development in agriculture is sluggish nowadays due to the attack of disease. Many farmers detect disease by their previous experience or some takes help from experts. But the experts usually judge the symptoms of disease with bare eyes. So there is the possibility of an inaccurate diagnosis of diseases having very large similarity in their symptoms. Any error during diagnosis of the disease sometimes may leads to wrong controlling and an excess use of pesticides [4]. Therefore, it is essential to move towards the new strategies for automatic diagnosis and controlling of disease. A number of varieties of pesticides are available to control disease and increase production but finding the most suitable and effective pesticide to control the infected disease is difficult and required experts advise which is time-consuming and expensive.

The presence of disease on the cotton plant is mainly reflected by symptoms on leaves. So there is need of an automatic, accurate and less expensive machine vision system for detection of disease from cotton leaf images and to suggest the proper pesticide as a solution. This paper focuses on detection of most commonly occurring diseases on cotton leaves and their controlling using IoT based android App. The disease is detected using SVM based regression algorithm written in python code in Raspberry Pi. In this paper four different sensors that is temperature, moisture, humidity and water sensors are used and interfaced with raspberry pi for soil quality monitoring. Two android app are used; one for displaying soil parameters and other for displaying disease information, for turning external devices such as sprinkler or motor ON/Off and for handling the movement of the whole system from one place to another to check soil parameters at different places. So this system is useful in large farm for accurate detection and controlling of disease.

II. Cotton Leaf Diseases

A. Bacterial Blight

Bacterial blight is bacterial disease mainly caused by the bacteria "Xanthomonas Campestris pv. Malvacearum" [4]. The symptoms of Bacterial blight starts as Dark green, water soaked angular spot of 1 to 5 mm on a leaf with red to brown border. At the beginning, these angular leaf spots appear as water-soaked areas which later on changes from dark brown to black color[6]. The spots on the lesion area of leaves may spread over the major veins of leaf and in later petioles and stems get infected and premature fall off of the leaves occur [12]. *Fig. 1(a),* shows Bacterial Blight infected leaf.

B. Alternaria

It is a fungal disease mainly caused by A. Alterneta or Alternaria macrospora [12]. The disease is most severe on the lower part of leaves as compared to the upper part and may get confused with the spots of bacterial leaf blight as the symptoms are nearly similar[11]. At the beginning, brown, gray-brown to tan colored small circular spots appears on leaves and vary from 1-10mm in size which later on become dry, dead with gray centers which crack and fall out [11]. Sometimes, old spots combine together and create irregular dead areas. *Fig. 1(b),* shows Alternaria infected leaf.

C. Cerespora

Cercospora is brought about by the Cercospora Gossypina [12]. The tainted leaf has red spot blemishes on the leaves which expand in distance across to around 2 cm. The spots are round or unpredictable in shape with yellowish, purple, dark brown or blackish fringes with white focuses [11]. The rakish leaf spot shows up because of the limitation of the lesion region by fine veins of the leaf. This malady influences more seasoned leaves of develop plants. *Fig. 1(c),* shows Cerespora infected leaf.

D. Grey Mildew

It is fungal disease generally caused by the Ramularia Areola Atk. in imperfect stage and Mycosphaerella areola in perfect stage [12]. At the initial stage, the infection appears as triangular, square or irregularly circular whitish spots of 3-4 mm size on leaves [11]. This disease primarily appears as irregular angular and pale translucent spots of diameter l-l0mm on older leaves of matured plant [8]. The infected regions are light to yellowish green in color on the upper surface. As disease infection increases, the small spots merge together forming bigger spots and the leaf tissues turn yellowish brown while whitish frosty growth appears on a lower surface but occasionally on the upper surface [8]. *Fig 1(d),* shows grey Mildew infected leaf.

E. Fusarium Wilt

It is the fungal ailment largely caused by Fusarium oxysporum [12]. The organism can assault cotton seedlings, but the sickness for the most part shows up when the plants are more matured[11]. The influenced plants first get to be distinctly darker green and hindered [4]. Latterly, the yellowing of the leaves and loss of foliage is observed. At first, the manifestations show up on lower leaves around the season of first blossom and the leaf edges shrink, turn yellow initially followed by brown with inward movement. *Fig 1(e)*, shows Fusarium Wilt infected leaf.

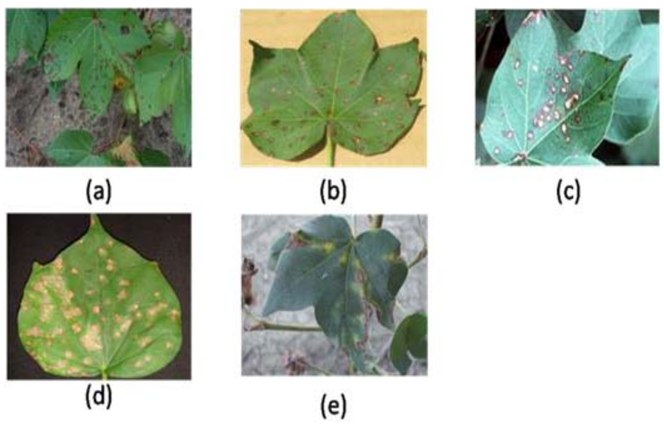

Fig. 1. Sample of diseased leaf (a) Bacterial Blight infected leaf[4] (b) Alternaria infected leaf (c) Cerespora infected leaf [12] (d) Grey Mildew infected leaf [12] (e) Fusarium Wilt infected leaf [2]

III. Literature survey

Very few National and international papers are surveyed and presented here.

Davoud Ashourloo, et. al.[1] proposed the machine learning techniques for Wheat Leaf Rust disease detection as well as evaluate the training sample size and influence of disease symptoms effects on the methods of predictions. This paper compares the performance of PLSR, ν-SVR, and GPR with the PRI and NBNDVI. The combinations of disease symptoms at each disease severity level results in complex spectra which declined the accuracies of PRI and NBNDVI while they do not have adverse impacts on PLSR, ν-SVR, and GPR performances. The GPR's performance using smaller training data set results in higher accuracy than other methods.

Rothe, P. R., et al.[4]proposed pattern recognition technique for the identification and classification of three cotton leaf disease i.e. Myrothecium, Alternaria, and Bacterial Blight. Segmentation of images is done by using active counter model and Hu's moments are extracted for the training of adaptive neuro-fuzzy inference system. Back propagation neural network(BPNN) is used for classification.

Viraj A. Gulhane, et. al.[5] used Principal Component Analysis(PCA) and Nearest Neighbourhood Classifier (KNN) for diagnosis of diseases on cotton leaf. After implementing PCA/KNN multi-variable techniques, statistical analysis of the data which is related to the Green (G) channel of RGB image is done. Here Green channel is taken into consideration for feature collection because disease or deficiencies of elements are reflected well by green channel.

Rothe, P. R. et al. [6] proposed the concept of automated feature extraction in which the RGB image is first captured and transferred to gray scale. The noise is removed by using LPF and Gaussian filter. K-means clustering followed by graph cut energy minimization operation is used for segmentation purpose. For color feature extraction, the segmented RGB image converted to YCbCr and DCT of YCbCr is obtained and zigzag scanning is used for achieving color layout descriptor of input image.

P. Revathi,et al.[9] proposed two phases to identify the lesion region of the disease. First Edge detection technique is used for segmentation and then image analysis and classification of diseases is done using the proposed HPCCDD Algorithm. This paper proposed RGB feature based techniques in which, the captured images are processed first and then color image segmentation is carried out to get disease spots. The edge features are extracted to identify the disease spots using Sobel and Canny filter.

IV. System architecture

The present system is used to detect and control the disease on cotton leaves along with soil quality monitoring. This system also helps to automatically ON/OFF the relay connected to external devices such as motor, sprinkler assembly according to need by using Android App. *Fig. 2,* depicts the block diagram of the system.

A. Block Diagram of the System

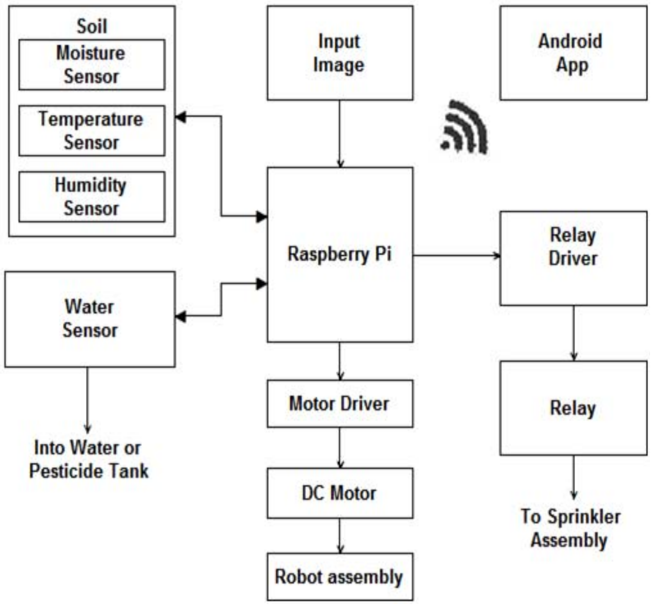

Fig. 2. Block diagram of system for cotton leaf disease detection and controlling

The present system consists of Raspberry Pi- model B which is the main part of the system used for interfacing purpose. Initially, the input image is selected. According to selected image disease is detected with its name and remedies by using python and to be displayed on the App. After disease detection farmers take the necessary action i.e turn ON/OFF the sprinkler assembly by using app to spray pesticides or fertilizers by mixing it in water. Relay driver and single pole double throw relay is used to control the ON/OFF of external devices. Farmer can also check the soil condition and water level in the tank with the help of sensor. Four different types of sensors are used for measuring soil condition and level of water or pesticide tank. These sensors include LM 35 temperature sensor, DHT-22 Humidity Sensor, Water sensor and moisture Sensor. All these sensors are interfaced with Raspberry Pi. The motor driver and DC motor are used for the movement of the overall system. The moving system help to monitor soil condition at different places.

B. Design flow of the Leaf Disease Detection

Fig. 3. depicts the design flow of cotton leaf disease detection. The disease detection is needed to be performed in step wise manner to get high accuracy. The main steps of disease detection are as:

1) *Image Acquisition*
2) *Pre-processing*
3) *Segmentation*
4) *Feature Extraction*
5) *Classification*

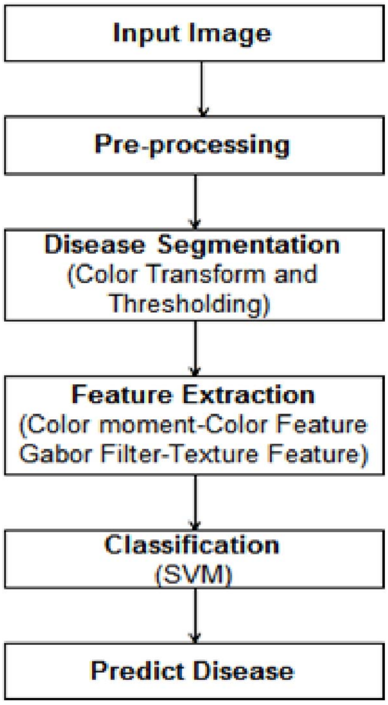

Fig. 3. Design flow for cotton leaf disease detection

1) Image Acquisition: Image acquisition is the principal phase which incorporates capturing of infected leaf images to build database[2]. The RGB color images of cotton leaves are captured utilizing Nikon digital camera in JPEG format from cotton cultivate in Buldhana district with required resolution which gives great quality pictures for disease detection. The database of 900 pictures are gathered. During this progression, noise is generated which is expelled by preprocessing. From database one image is taken and processed further as shown in *Fig. 4. Fig 4(a),* shows infected image of cotton leaf.To change the default, adjust the template as follows.

2) Pre-processing: The input image is pre-processed to enhance its quality by expelling the unwanted noise and to encourage the further process. The aim of pre-processing is improvement of an image data that suppress unwanted distortions and enhances some image features important for further processing. The preprocessing methods incorporates image enhancement, color conversion, resizing and filtering of an images. In present system resizing and median filter is adopted for pre-processing to expel noise and to get good quality image. First the image is read and then resized into 250×250 pixel and then median filter is applied. Median filter is used as it is more accurate as compared to mean filter and preserve edges while removing noise. *Fig. 4(b),* depicts the image obtained by pre-processing the captured image in fig 4(a).

3) Segmentation: Segmentation is an important step in leaf diseases detection which is used to extract the lesion region from images [2]. The Segmentation separates the region of connected pixel having similar properties by framing limits between these regions. In present system, color transformation and thresholding is used to extract the region of interest (ROI) from the image. First the RGB color format image is transformed to YCbCr color format. After color conversion, Bi-Level thresholding is applied. In bi-level thresholding two threshold range such as upper and lower threshold is defined for each of three color plane i.e. Y, Cb, Cr and pixel between these ranges are considered as diseased part. After applying bi-level thresholding, we get logical black and white image which is needed to be further processed to extract region of interest (ROI). *Fig. 4(c),* depicts the YCbCr color transformed image while *Fig. (d),* shows logical black and white image.

4) Color Mapping: In color mapping, the logical black and white image is converted to RGB color format which is a masked image. Then bit wise operation is performed to get the RGB masked image. We are interested only in diseased part so this RGB masked image is converted to grey image. *Fig 4(e),* shows the RGB masked image while *Fig 4(f),* shows the gray image. In this grey image white color shows diseased part which is our Region of Interest (ROI).

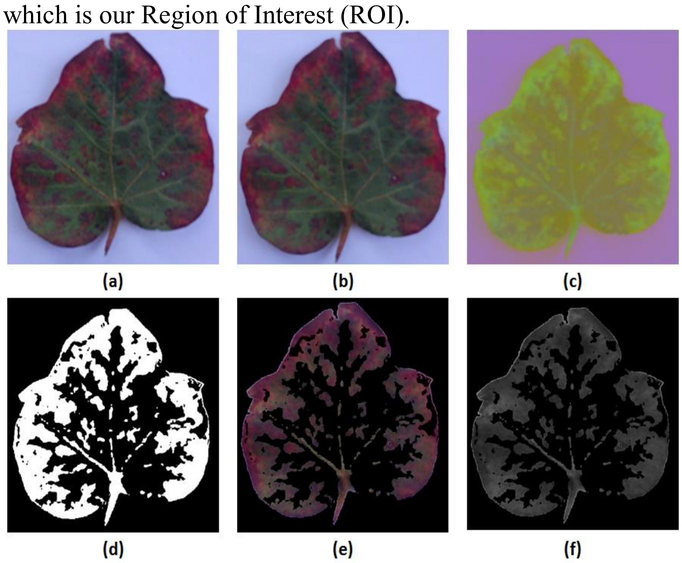

Fig. 4. Experimental results of disease detection steps: (a) RGB input image of infected cotton leaf (b) Pre-processed image (c) YCbCr color transformed image. (d) Logical Black and White image (e) RGB masked image (f) Grey Level Image

5) Feature Extraction: Features extraction is the next important step after segmentation to perform. Feature extraction aims at the extraction of the relevant information that characterizes different classes [8]. The goal of feature extraction is to extract a set of features representing each character, which maximizes the recognition rate with the least number of elements. The diseased area separated by segmentation is the region of interest (ROI) in feature extraction in order to extract different features that used to identify the disease. Total eight color and texture features are extracted in present system using partial least square regression(PLSR). PLSR finds a few components in order to explain most of the variations in both predictors and responses.

a) Color Features: Color features are extracted by various methodologies such as color correlation, color histogram, color descriptor and color moment. In the present work color moment is used for color feature extraction. In color moment, Mean and standard deviation are embraced. Therefore, an image is characterized by 6 moments as 2 color moment for each 3 color channels (i.e. 3*2=6). Here we define the i-th color channel at j-th image pixel as Pij. Mean and standard deviation can be defined as:

Moment 1- Mean: It represents the average color value of the image.

$$Mean = E_i = \sum_{j=1}^{N} P_{ij} \tag{1}$$

Moment 2- Standard Deviation: Std. Deviation represents the square root of the variance of the distribution.

$$Std.Deviation = \sqrt{\frac{1}{N}\sum_{j=1}^{N}\left(P_{ij} - E_i\right)^2} \tag{2}$$

b) Texture features: In present work, two texture features are extracted using 2D Gabor filter. The Gabor texture feature include mean and std. deviation of the magnitude of Gabor wavelet transform coefficient. Mean gives the average value and std. deviation decides the deviation of data from center of Gaussian filters. In present system, Gabor filter with different parameters such as frequency, angle in 16 orientation with 10 different frequencies is considered.

6) Classification: This is the last step of disease detection in which classifier is used to recognize the kind of the leaf disease. Classification deals with matching the given data vectors with one of the trained data of different classes. In machine learning, there are various types of classifier available for classification. To get high accuracy, SVM is trained and tested with different kernels and come to know that Gaussian kernel gives high accuracy. In present system SVM based regression technique with non-linear Gaussian kernel is used for classification of the diseases on cotton leaves. The SVM based regression finds the nonlinear relationship between input vectors and response variables by finding the best hyperplane. The best hyperplane is that one which is at maximum distance from test vectors.

V. EXPERIMENTAL RESULTS

As discussed, the main aim of the present work is to detect and control diseases on cotton leaves. The secondary aim is to monitor the soil quality. For disease detection, we used 900 images of cotton leaves. Out of this 629 images are trained and 271 images are used for testing. Table 1 shows the accuracy for individual cotton leaf disease detection. The fig.5 shows the graphical representation of table 1. *Fig.6,* shows the implementation of system showing raspberry Pi, Relay and sensors. *Fig.7(a),* shows the result of android app displaying cotton leaf disease name with its remedies. It also shows forward, reverse, left and right movement and stop options along with motor ON/OFF to ON or OFF the relay. *Fig.7(b),* shows the result of android app displaying different sensors reading.

Table 1: Accuracy of Cotton leaf Disease detection

Name of Disease	Correctly Classified (CC)	Incorrectly Classified (IC)	Accuracy of CC
Bacterial Blight	67	11	85.89%
Alternaria	55	10	84.61%
Cerespora	39	08	82.97%
Grey Mildew	31	06	83.78%
Fusarium Wilt	28	06	82.35%
Healthy leaf	08	02	80%

The overall accuracy of proposed system for disease detection is given by equation 1 and it is 83.26%.

$$Overall\ Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (3)$$

Where TP- True Positive, TN- True Negative, FP- False positive, FN- False Negative

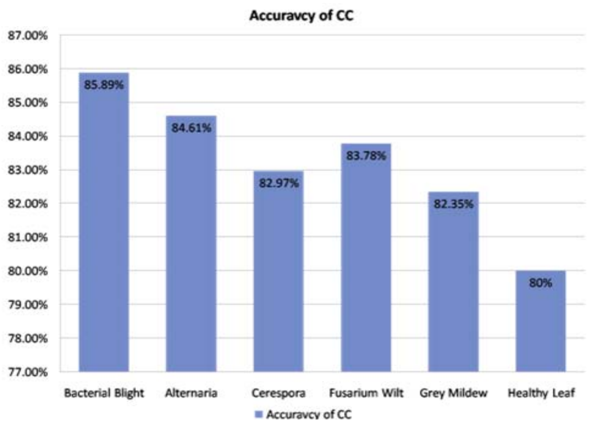

Fig. 5. Graphical representation of table 1

Fig. 6. Overall system implementation

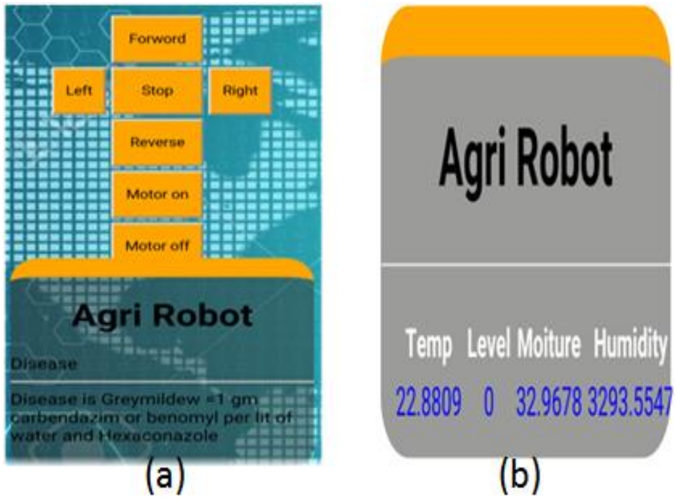

Fig. 7. (a) Android app displaying cotton leaf disease information (b) Android app displaying different sensors reading.

VI. Conclusion

This paper presents the Support Vector Machine based regression technique for detection of five cotton leaf disease. The pesticides as a remedy is suggested to the farmer for infected disease to control it. The android app is developed to display disease and sensor information along with the ON/OFF of the relay. The app also handle the movement of the whole system from one place to another. So by using present system, the farmers can automatically detect the disease and know the remedies to control that disease. The farmer can move the system from one place to another place to check the soil condition at different locations with the help of sensor and can change soil condition by turning motor ON/OFF with relay. All these processes are done by using android App which saves human hard work in large field area. The use of Raspberry pi make this system cost effective and independent. The present system gives accuracy of 83.26% for disease detection and proves its effectiveness to the farmers for cotton leaf disease detection and controlling by improving the crop production.

References

[1] Davoud Ashourloo, Hossein Aghighi, Ali Akbar Matkan, Mohammad Reza Mobasheri and Amir Moeini Rad, “An Investigation Into Machine Learning Regression Techniques for the Leaf Rust Disease Detection Using Hyperspectral Measurement” *IEEE Journal Of Selected Topics In Applied Earth Observations And Remote Sensing*, pp.1-7, May 26, 2016.

[2] Youssef Es-saady, Ismail El Massi, Mostafa El Yassa, Driss Mammass and Abdeslam Benazoun, “Automatic recognition of plant leaves diseases based on serial combination of two SVM classifiers” *2nd International Conference on Electrical and Information Technologies(ICEIT), IEEE*, 2016.

[3] Sonali Dash, K.Chiranjeevi, Dr.U.R.Jena and akula.Trinadh, “Comparative study of image texture classification technique”,*International Conference on Electrical, Electronics, Signals, Communication and Optimization IEEE*, 2015.

[4] P. R. Rothe and R. V. Kshirsagar, “Cotton Leaf Disease Identification using Pattern Recognition Techniques”, *International Conference on Pervasive Computing (ICPC) IEEE*, pp. 1-6, 2015.

[5] Viraj A. Gulhane and MaheshkumaR. H. Kolekar, “Diagnosis Of Diseases On Cotton Leaves Using Principal Component Analysis Classifier” *Annual IEEE India Conference (INDICON),IEEE*,2014.

[6] P. R. Rothe, Dr. R. V. Kshirsagar, "Automated Extraction of Digital Images Features of three kinds of Cotton Leaf Diseases", *International Conference on Electronics, Communication and Computational Engineering IEEE, 2014.*

[7] Sanjeev S. Sannakki, Vijay S Rajpurohit, V. B. Nargund and PallaviKulkarni, “Diagnosis and Classification of Grape Leaf Diseases using Neural Networks”, *International Conference on Computing Communications and Networking Technologies IEEE, 2013.*

[8] P. Revathi and M. Hemalatha, “Advance Computing Enrichment Evaluation of Cotton Leaf Spot Disease Detection Using Image Edge detection”, *IEEE*, 28th July 2012.

[9] P. Revathi and M. Hemalatha “Classification Of Cotton Leaf Spot Disease Using Image Processing Edge Detection Technique”, *International Conference on Emerging Trends in Science, Engineering and Technology IEEE*, PP. 169-173, 2012.

[10] Cotton disease management. http://www.ikisan.com/mhcottondiseasemanagement.html

[11] Plant village Cotton. https://www.plantvillage.org/en/topics/cotton

[12] Texas plant disease handbook. http://plantdiseasehandbook.tamu.edu/industryspecialty/fiber-oil-specialty/cotton/