ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 6, June 2015

Leaves Classification Using SVM and Neural Network for Disease Detection

Bhushan R. Adsule, Jaya M. Bhattad

M.Tech Student, Department of Electronics and Telecommunication, Priyadarshini College of Engineering, Nagpur, Maharashtra, India

Assistant Professor, Department of Electronics and Telecommunication, Priyadarshini College of Engineering Nagpur, Maharashtra, India

ABSTRACT: The process to classify cotton, orange and Lemon leaf diseases using neural network analysis and support vector machine. A few of infected leaf samples were collected and they were captured using a digital camera with specific calibration procedure under controlled environment. The classification on the leaves diseases is based on colour feature extraction from RGB colour model where the RGB pixel colour indices have been extracted from the identified Regions of Interest (ROI). The proposed automated classifications model involved the process of diseases classification using Support Vector Machine and Neural Network Pattern Recognition Toolbox in MATLAB. The proposed techniques based on performance indices results are promising with accuracy.

KEYWORDS*:* Leaf Diseases; RGB; Neural Network Pattern Recognition; Support Vector Machine; Image Segmentation.

I. INTRODUCTION

India is a large country that rich with verities plants and fruits. India is second largest farm output producer in world. Climate of India is hot in area of Maharashtra in Vidarbha. The agriculture is most important business and income source of India. Many acres of area in India covered under agriculture. So many farmer's life depend on agriculture but cultivators face so many problems because of several diseases. Diseases directly effect on quality and quantity of crop. So it is more challenging task for cultivators to prevent crop from diseases. By using plant leaves we will try detect the diseases of cotton, orange and Lemon plant by applying some image processing techniques to identify diseases. From this technique we will try to capture disease at early stage and prevent them. This is more useful project for farmer and my little contribution for them. Some of the popular Cotton, Orange leaf diseases in India are Anthracnose, Downey Mildew and Powdery Mildew.

The main diseases that cause serious effect to the production of orange and cotton are *Downey Mildew* and *Anthracnose*. Between these two diseases, *Downey Mildew* is dominant compare to *Anthracnose*. *Downey Mildew* is caused by fungus called *Pseudoperonospora cubensis.* This disease is most extreme during rain and low temperature. Misshapen yellowish to brown Patches, often indefinite in outline, appear on upper leaf surface near the tip. Brown patches later become more distinct on both sides of the leaves[2]. Likewise, *Anthracnose* is caused by fungus *Colletotrichum orbiculare.* The symptoms appear first on tip of leaves as brown-black patches. These lesions are also visible on the underside of leaves. During moist weather, orange-pink masses of spores develop in center of larger leaf. During dry weather the spore masses turn to gray [5]. These diseases will affect the sweetness, size, and freshness of fruit.

II. LITEARTURE SURVEY

The related work about the image for pattern detection applied by the author. It is difficult extract the leaf captured samples from sunlight background, having poor imaging conditions and different plant material. To settle the problem, color image segmentation approach consisting of fuzzy C-means clustering, and morphological operation was

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 6, June 2015

proposed by author[3]. Due to low contrast in the images captured from natural light or sunlight low range of pixel distribute in the colour space, both contrast adjusting algorithm and decorrelation stretch transform were for image process. Then morphological filtering performed for removing all the noises in image. Experimental results showed that the proposed approach is effective for extracting plant leaf images from complicated background and it can meet the demand of plant real time monitoring [3]. The purpose of the study is to recognise and fragment the diseased-pattern section that consist of colour and texture, apart from the background which is also known as region of interest (ROI)[2]. Hence, a comparative study related to segmentation of Elaeis Guineensis (oil palm) leaf images extracted from the oil palm will be evaluated and validated. Three different techniques of segmentation are investigated specifically the local threshold and global threshold. Subsequent, the segmentation algorithms have been developed to be proficient to perform segmentation process using the leaf images that were visible to varying radiance. Initial findings showed that global threshold is the best segmentation based on the tested images [2]**.** This author's work about segmentation and disease pattern detection and effective experimental results shows.

III. METHODOLOGY

A. Data Collection

A total of 200 leaf samples were taken from a Cotton Orange and Lemon nursery in Nagpur. The leaf samples were collected carefully to make sure that the symptoms are related to the disease. The leaves samples were then put into the black plastic bags in order to preserve the freshness. The calibration of the room lighting need to be done prior to photo session in order to make sure the consistency of lux intensity in the studio is under controlled environment. The photo session was taken placed at the Image Capturing Studio Room (ICS Room) at Advanced Signal Processing (ASP) Research Lab, Faculty of Electronics and Telecommunication, PCE. Fig. 1 below shows the samples of Cotton leaf showing normality and abnormality.

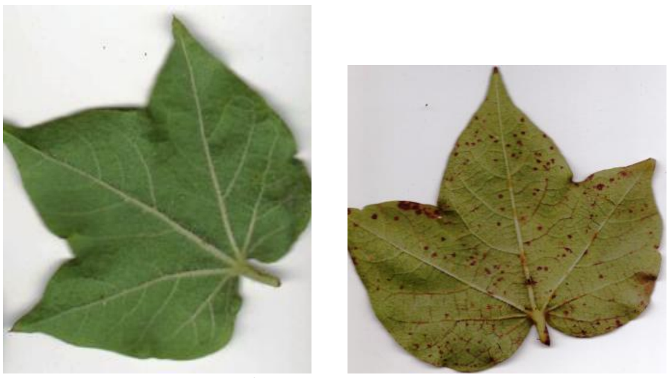

Fig.1 Cotton leaf samples Normal and Abnormal

B. Data Capturing Setup

The basic digital colour image in (RGB) were obtained using Nikon D80 digital camera with image resolution of 3872 x 2592. For the camera settings, the International Standards Organization (ISO) was set at 800 to gain the sensitivity of image sensor, the shutter is 50, the aperture is F8 and zoom is 35mm. For capturing the image camera was placed 11.5 inches above from the leaf sample at 90 degrees angle. The lighting source was provided by Further details on image acquisition setup are illustrated in Fig. 2.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 6, June 2015

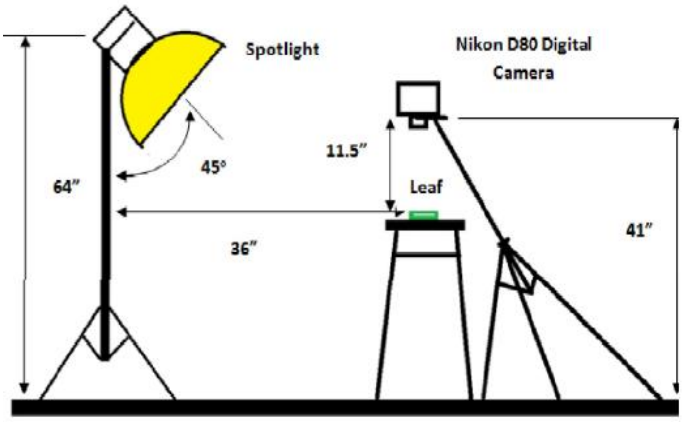

Fig.2 Image acquisition setup.

C. Image Processing

Initially, the Region of Interest (ROI) needs to be identified from each infected leaf samples based on its RGB colour component. Then, it is ready to crop out and resized into a dimension of 15 x 15 pixel area [2]. Later, all cropped data will be transferred to further analysis.

The number of samples obtained from the cropping process was 100 and 91 for Downey Mildew and Anthracnose respectively.

D. Preprocessing

Conventionally, a colour image of size M×N can be defined as 3 sequences of pixel level:-

$$r(m, n),\ 1 \le m \le M\ ;\ 1 \le n \le N \quad (1)$$
$$g(m,n),\ 1 \le m \le M\ ;\ 1 \le n \le N \quad (2)$$
$$b(m,n),\ 1 \le m \le M\ ;\ 1 \le n \le N \quad (3)$$

Where RGB colour component represents each pixel of the image[1]. The independent colour components are generally represented by 8-bit which means that each pixel would be an integer in the interval [0,255]. In any processing of 8-bit images, the integer restriction is unrestrained and a floating point representation in order to minimize quantization effects in image [1].

E. Median Filtering

Preliminary step in the process was the preprocessing of the images in order to reducing noise and facilitating image segmentation by using median filtering [1,7]. The imaging technique may be noisy in terms of small white ellipse lines or dots. This artifact can be considered as impulsive noise and may thus be reduced using a median filter given by:

$$P_{med}(m,,n) = median\{P(m-k, n-1) \mid \ -\frac{N_{med}-1}{2} \le k,$$
$$l \le \frac{N_{med}-1}{2} \wedge 1 \le m-k \le m \wedge 1 \le n-1 \le N\ \} \quad (4)$$

Where Nmed is odd2 and indicates the size of the two dimensional median filter. P represents all the three color components and only square median filter kernel was considered [7].

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 6, June 2015

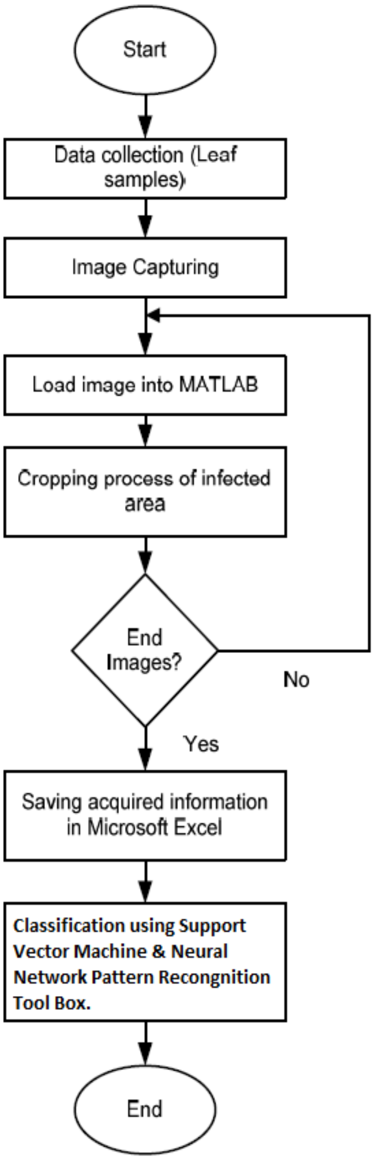

Fig.3 Flowchart of identification leaf diseases

F. Image Segmentation

The segmentation technique involves the image partition such that the diseased region converts the grayscale image into binary (black and white) [3]. The image of leaf is marked as foreground or ROI while the remaining area is assigned as background. In order to segment the leaf image, threshold value needs to select appropriately from histogram distribution of gray scale image. Generally, the common distribution of histogram produces bimodal peak intensity to represent the ROI and background, thus a good threshold will generate one value of intensity [7]. The basic segmentation approaches are proposed towards leaf image, namely global, Basic global thresholding

G. Global Thresholding

Global threshold selects the value of threshold by separating the leaf image peak intensity into two classes namely background and foreground while minimizing the overlap between clusters. The value is automatically selected by minimizing the weighted within class variance and assuming maximizing of the class variance [2]. This can be represented by:

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 6, June 2015

$$g(x,y) = \begin{cases} 1\, if \;\; f(x,y) > T \\ 0\, if \;\; f(x,y) \le T \end{cases} \quad (5)$$

Where: T = threshold value, 0.5 (μ1 + μ2)

H. Support Vector Machine Toolbox

It is binary liner classifier supervised learning models with associated learning algorithms that analyses data and recognize patterns, used for classification and analysis. It has a simple geometrical interpretation in a high-dimensional feature space that is nonlinearly related to input space. By using kernels all computations keep simple. It contains ANN, RBF and Polynomial classifiers as special cases.

I. Neural Network Pattern Recognition Toolbox

This work develops a classification system for leaf diseases using Neural Network Pattern Recognition Toolbox (nprtool) in MATLAB. Learning vector quantization (LVQ) is a method for training competitive layers in a supervised manner. If two input data vectors are very identical, the competitive layer probably will put them in the same class.

III. RESULTS AND DISCUSSION

A .Support Vector Machine

The first analysis is made by using Support Vector Machine Tool in order to enhance Classification process. The accuracy of testing normal leaves are Orange (63%), Cotton (71.15%) and Lemon (71.3%) and Abnormal are Orange (69.7%), Cotton (70.59%) and Lemon (70.29%) calculated from the classification. The overall accuracy of normal leaf is 68.81% and for abnormal leaf is 70.21%.

TABLEI. OVERALL PERFORMANCE FOR SVM BOTH NORMAL ABNORMAL

SVM	Normal			Abnormal		
Orange/Cotton/ Lemon	O	C	L	O	C	L
Accuracy of Testing (%)	63%	72.15%	71.3%	69.7%	70.59%	70.29%
Overall (%)	68.81%			70.21%		

B. Neural Network Pattern Recognition

The second analysis is made by using Neural Network Pattern Recognition Tool in order to enhance the previous analysis with the best 250 neurons in its hidden layer. The plots of confusion matrix for each training, validation, testing and overall performance in Neural Network were shown in Figure.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 6, June 2015

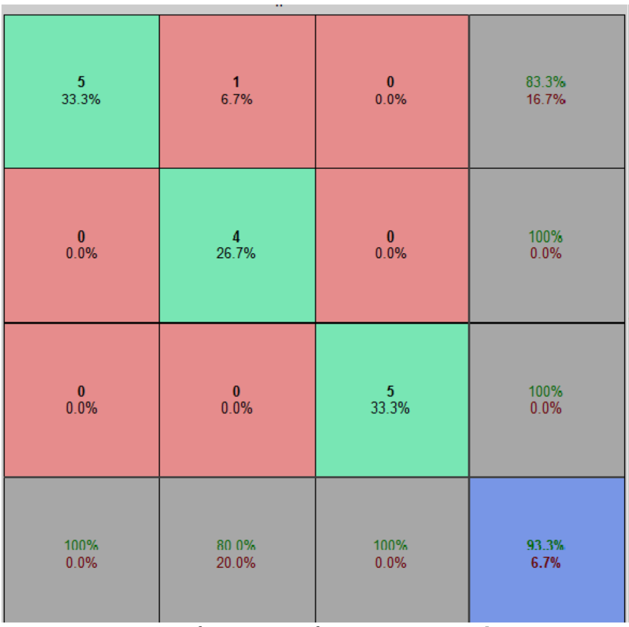

Fig.4 Confusion matrix for Orange training data.

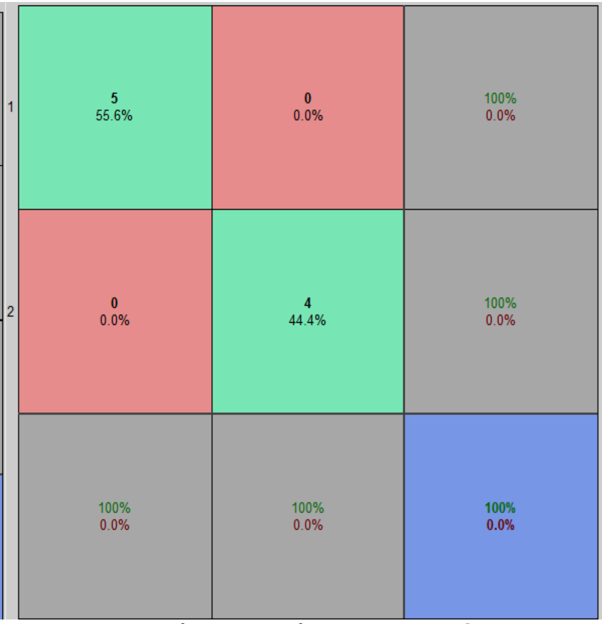

Fig.5 Confusion matrix for Lemon training data

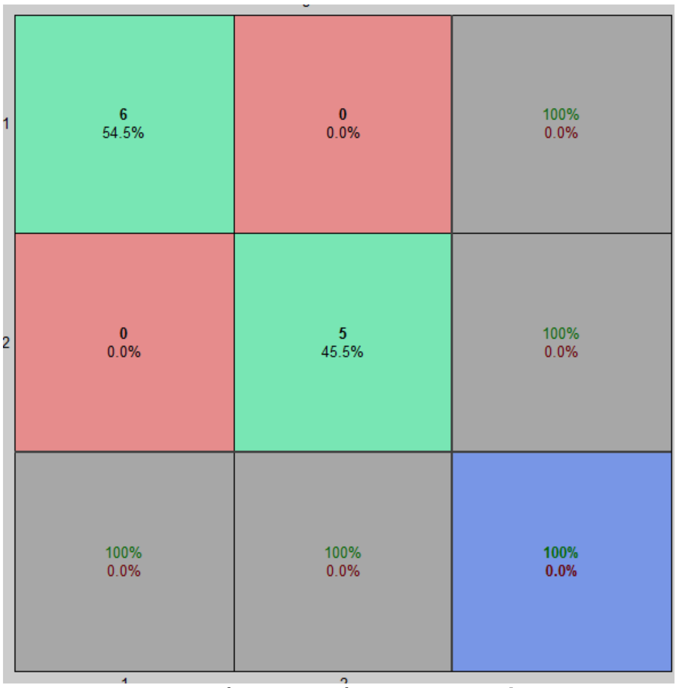

Fig.6 Confusion matrix for Cotton training data

The training confusion matrix plot is consists of True Positive Rate (TPR) represent by Green box, True Negative Rate (TNR) represent by Red box and overall accuracy (Blue box). While row number 1 for orange leaf disease Based on Fig. 5 the training data is set to be 25% (30 samples). From the confusion plot it is observed that the TPR and TNR for Orange leaf disease normality and abnormality are 33.3% and 6.7% respectively. The overall accuracy for training is 93.3%.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 6, June 2015

The training confusion matrix plot is consists of True Positive Rate (TPR) represent by Green box, True Negative Rate (TNR) represent by Red box and overall accuracy (Blue box). While row number 1 for Lemon leaf disease Based on Fig. 6 the training data is set to be 25% (30 samples). From the confusion plot it is observed that the TPR and TNR for Orange leaf disease normality and abnormality are 55.5% and 0.0% respectively. The overall accuracy for training is 100%. While row number 1 for Cotton leaf disease Based on Fig. 6 the training data is set to be 25% (30 samples). From Fig.7 the confusion plot it is observed that the TPR and TNR for Orange leaf disease normality and abnormality are 54.54% and 0.0% respectively. The overall accuracy for training is 100%.

The confusion matrix for 100% (45 samples) of cotton leaf test data. From the plot it is observed that the TPR and TNR for the cotton leaf Abnormality are 76.41% and 23.59% respectively while the Normality achieved TPR (93.33%) and TNR (6.67%). It is indicated that the accuracy of testing is 94%.

The confusion matrix for 100% (34 samples) of Orange leaf test data. From the plot it is observed that the TPR and TNR for the Orange leaf normality are (90.90%) and (9.09%) respectively while the abnormality achieved TPR (96%) and TNR (4%). It is indicated that the accuracy of testing is 93.45%.

The confusion matrix for 100% (14 samples) of Lemon leaf test data. From the plot it is observed that the TPR and TNR for the Orange leaf Abnormality are (90%) and (10%) respectively while the normality achieved TPR (95%) and TNR (4%). It is indicated that the accuracy of testing is 93.45%.

The overall average accuracy of neural network is 96.27% calculated.

TABLEII. OVERALL PERFORMANCE FOR BOTH NORMAL ABNORMAL

Neural Network	Normal			Abnormal		
Orange/ Cotton/Lemon	**O**	**C**	**L**	**O**	**C**	**L**
Colour Pixel Component	RGB mean			RGB mean		
Accuracy of Training (%)	100%	100%	100%	100%	100%	100%
Accuracy of Testing (%)	90.90%	90%	95%	96%	93.33%	90%
Overall (%)	95.45%	95%	97.5%	98%	96.67	95%

IV. CONCLUSION

This paper describes the analysis of Cotton, Orange and lemon leaf classification by using image processing technique with respect to its mean value of RGB colour component. There are two types of leaf diseases used in this work named Downey Mildew and Anthracnose. Each group consists of images captured under specific requirement. The cropped data from each leaves images then were analysed and then tested their performance by using Support Vector Machine and Neural Network Pattern Recognition Toolbox [8]. The performance is measured in terms of accuracy of the system.

V. FUTURE RECOMMENDATION

In order to improve the effectiveness of this classification system for the leaf diseases, it is recommended to use a high pixel of digital camera to get the best images. Also recommended to increase the number of data for the training and testing to get the best result. In addition, the lighting setup must be in proper position because it also can affect the image captured. Besides that, the other color model can be used as the input in order to increase the efficiency such as Cyan, Magenta, Yellow (CMY), Hue Saturation Value (HSV) and Hue Saturation Lightness (HSL).

VI. ACKNOWLEDGMENT

The author would like to thank to Dr. Sidharth M. Wasnik (Principle Scientist) at Central Institute Of Cotton Research, Nagpur for sharing their knowledge and providing help for research.

REFERENCES

1. "Classification of Watermelon Leaf Diseases Using Neural Network Analysis" 2013 IEEE Business Engineering and Industrial Applications Colloquium (BEIAC) Suhaili Beeran Kutty, Noor Ezan Abdullah, Dr. Hadzli Hashim, A'zraa Afhzan Ab Rahim, Faculty of Electrical Engineering University Technology, Mara Shah Alam, Selangor
2. Muhammad Asraf Hairuddin, Nooritawati Md Tahir, Shah Rizam Shah Baki and Nur Dalila Khirul Ashar "Elaeis Guineensis Leaf Image Segmentation:
3. "A Comparative Study and Analysis" 19 - 20 Aug. 2013, Shah Alam, Malaysia.
4. Lin Kaiyan, Wu Juhui, Chen Jie "A Real Time Image segmentation Approach For Crop Leaf" 2013 fifth conference of Measuring Technology & Mechatronics Automation, Shanghai, China.
5. Abdullah."Classification of Rubber Tree Leaf Diseases using Multilayer Perceptron Neural Network" in The 5th Student Conference on Research and Development (SCOReD 2007*)*.December 2007: Kuala Lumpur.
6. http://aggiehorticulture.tamu.edu/vegetable/watermelon/foliar.text.html view on January 2012 Volume-3, Issue-4, September 2012.
7. Central Institute for Cotton Research, Nagpur. www.cicr.org.in/
8. A. K. Jain, Fundamentals of Digital Image Processing: Prentice Hall, 1989.
9. MATLAB, *Neural Network Pattern Recognition Toolbox and Support Vector Machine 2014.*

.

BIOGRAPHY

Mr. Bhushan R. Adsule is a M.Tech research student in the Electronics And Telecommunication Department, Priyadarshini College Of Engineering, Nagpur. He received Bachelor Of Enginnering (B.E) degree in 2011 from PRMIT&R, Amaravati, MS, India. His research interests are Communication (Image Processing), Wireless Communication etc.