ISSN 2229-5518

Detection and Grading of diseases in Banana leaves using Machine Learning

Vipinadas.M.J, A.Thamizharasi

Abstract— The diseases on the banana are a crucial issue which makes the sharp decrease in the production of banana. These require careful diagnosis and timely handling to protect the crops from heavy losses. Now a day's crop faces many diseases. The naked eye observation of experts is the main approach adopted in practice for detection and identification of banana leaf diseases. But, this needs continuous monitoring of experts which might be prohibitively expensive in large farms. So automatic detection of banana plant diseases are an important research topic as it may prove benefits in monitoring large field of crops, and thus automatically detect diseases from symptoms that appear on plant leaves. For detecting diseases on the banana leaf early and accurately we can use image processing technique. This paper provides a various methods to detect plant diseases using image processing technique. The proposed system is also a well organized module that identifies the Blacksigatoka disease and Panama wilt disease on banana leaf. The disease grading has been done using ANFIS classifier. Finally, classifiers comparison has been performed using confusion matrix.

Index Terms—banana leaf, disease identification, SVM, ANFIS, sigatoka, Panamawilt, classification etc

1 INTRODUCTION

Agriculture is the prime factor for food production and it is an important occupation for most of the Indian families. In India, agriculture gives about sixteen percent (16%) of total GDP and ten percent (10%) of total exports. If any complication occurs in the area of agriculture, it will directly or indirectly affect our economy and the population.

The management of crops requires close monitoring especially for the identification of diseases that can affect production and economy of a country. The Machine Learning can be used to detect the diseases in crops. There are many things that can cause different diseases to the plants, which leads to damages in agricultural field and completely the economy of the country. If plant diseases are detected effectively and managed accurately, then we can avoids the losses in crop fields. So to make the agricultural fields and the economy of the country stronger, rapid and accurate disease detection system is needed.

Banana (Musa spp) is a vital food in many countries of the humid tropics. It is among the most important carbohydrate sources in the diet of people in these regions. It needs low labor requirement and relatively high-energy output make plantain a suitable vital for areas where labor shortage is usually the main constraint to production. Despite the fact that crop is grown in diverse agro climatic conditions of the country, its production is continuously being hindered by a number of diseases inflicting yield losses of staggering dimension both in quantity as well as quality aspects.. The banana cultivation may be affected by certain implicit factors like Nutrients imbalance and Bacterial or fungal attacks and explicit factors like floods, droughts etc.

The banana leaves are affected by many diseases such as Blacksigatoka, Panamawilt and Mosaic disease. In Blacksigatoka, a pale yellow/dark or brown/black streaks are present parallel to the vein of the leaf. These streaks enlarge to form characteristic spots having dark brownish to black color, linear-oblong areas with ill-defined border, often surrounded by yellow halo as shown in figure 1.

- *Vipinadas.M.J is currently pursuing M.tec in computer science and engineering in Mohandas College of Engineering, Kerala University, India. E-mail: vipinamohandas@gmail.com*
- *A.Thamizharasi,Assistant professor is Mohandas College of Engineering, Kerala University, India. E-mail: radhatamil@gmail.com*

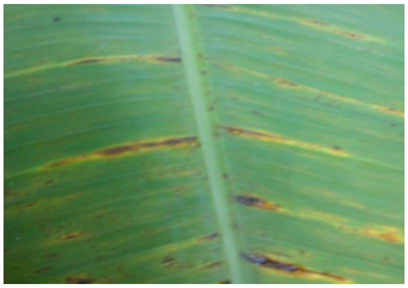

Fig. 1. Black Sigatoka disease on banana leaf

The Panamawilt caused by Fusarium oxysporum f.sp.cubens fungus. It infects all stages of crop growth. Its main symptom is characteristic yellowing of older leaves accompanied by vascular discoloration in the outer leaf sheath. If the leaves wilt, then the petiole buckles and the leaf hangs between the pseudo stem and the middle of lamina, while the leaf is still green as shown in figure 2.

Fig. 2.Panama disease in banana leaf

2 LITERATURE REVIEW

Sanjeev S Sannakki, Vijay S Rajpurohit, Nargund, Arun Kumar R and Prema S Yallur [1] proposed a new approach which converts the RGB color space into L*a*b color space. The K-means clustering is used to segment the diseased portion from the total leaf area. The fuzzy logic is used to calculate the percentage of infection in the diseased leaf.

Recently, Sabah Bashir, Navdeep Sharma [2] introduced CCM (color co-occurrence matrices) method and texture features are used to recognize and Classify different diseases in Malus Domestica [16]. The K-means clustering algorithm is used to classify pixels based on a set of features into K number of classes. The classification has been done by minimizing the sum of squares of distances among the objects and the cluster or class centroid. The Bayers classifier is used for the classification.

3 PROPOSED SYSTEM

The proposed system presents an efficient method for image pattern classification in Banana leaf disease identification. First, the blacksigatoka diseases on banana leaf are detected using svm and anfis classifier. The classifiers comparison is performed using svm and anfis classifiers. Then diseases on the banana leaves are graded using anfis. Finally, diseases are classified as blacksigatoka disease or panamawilt disease using multiple level svm. The proposed system is shown in figure 3.

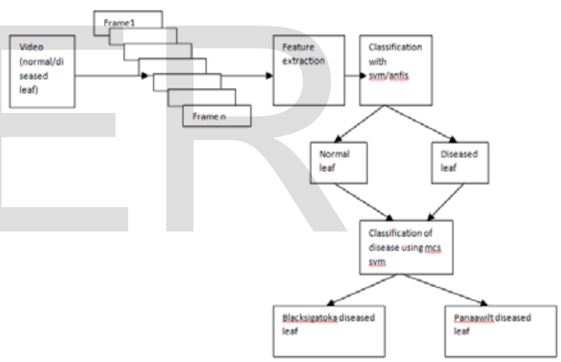

Fig. 3.Proposed system

3.1 Identification of disease in banana leaf using Support Vector Machine or ANFIS

The aim of the work is to check whether the banana leaf is diseased or not. The first step is to detect whether the leaf is infected by Sigatoka leaf spot disease or not. The input image is the RGB color image of banana leaf. The RGB image is first converted into YCbCr color space and the gray scale image is obtained by using 'Y' component. The next step is to apply adaptive contrast map to enhance the leaf diseased portions. This image is then converted into binary image using a threshold. The threshold value is set to 0.18. Then morphological operations like dilation, filing holes are applied. The resultant image obtained shows white pixels in the diseased area and black pixels in the normal portions [7].

The color, texture and shape features are extracted for classification. The color features extracted are skewness, kurtosis and mean. Shape features include solidity, extent, minor axis length and eccentricity. The Gray-Level Co-occurrence Matrix (GLCM)[3] is used to extract the texture features [7]. The texture features extracted are contrast, correlation, energy, and homogeneity. These 17 features are classified using Support Vector Machine (SVM)[5] classifier or ANFIS[6]. From the classification output determines whether the leaf is diseased or not. The figure.4 shows the block diagram for the steps in disease detection.

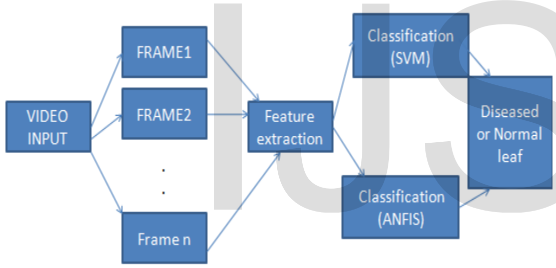

Fig. 4.Disease detection system

3.1.1 Adaptive Neuro-Fuzzy Inference System

Adaptive Neuro-Fuzzy Inference System (ANFIS) is a fuzzy toolbox .It integrates Neural and fuzzy logic. It combines the learning capabilities of neural networks and explanation capabilities of fuzzy system. As a result neural networks become more transparent and fuzzy logic capable of learning.

ANFIS network control systems (or neuro-fuzzy systems) represent a hybrid platform for solving actual complex problems that require the use of intelligent systems and are a viable alternative to the conventional model based control schemes. They allow dealing effectively with the common issues of uncertainty and unknown variations in plant parameters and structure, hence improving robustness of the control system.

A first-order Sugeno-style FIS model is a system that manages the process of mapping from a given crisp input to a crisp output, using fuzzy set theory.

If $x1$ is A and $x2$ is B then y=$f(x1, x2)$ (1)

Where A and B are fuzzy sets and y=$f(x1, x2)$ is a function associated to the fact of $x1$ and $x2$ pertaining to A and B respectively. $f(x1, x2)$ will usually be a polynomial, in what is then called the first-order Sugeno fuzzy model, in contrast with the zero-order Sugeno fuzzy model if f(x1, x2) were constant. Figure 5 shows the fuzzy reasoning procedure for a first-order Sugeno fuzzy model:

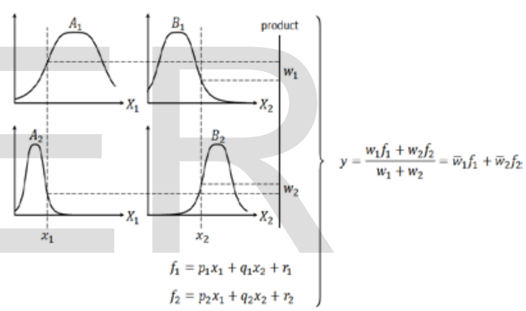

Fig. 5. The Sugeno fuzzy model

3.1.2 Support Vector Machine

The support vector machine is a binary classifier, which classifies the data into two classes the positive class and the negative class. The classification has been performed by using a separating plane called hyper plane. The figure 6 shows the diagrammatic representation of support vector machine [8].

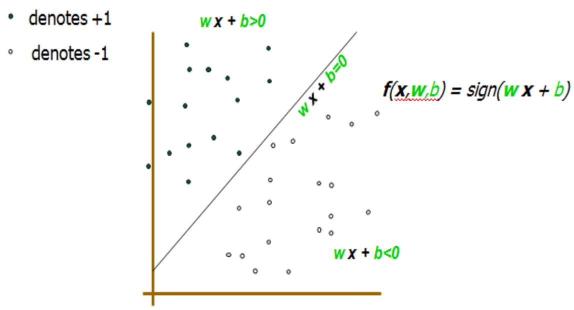

Fig. 6. The Support Vector Machine model

Here positive class represents the diseased leaf features and negative class represents the normal leaf features.

3.2 Classifier comparison

The ANFIS and SVM classifiers are used for classifier comparison. The confusion matrix has been used for classifier comparison.

3.2.1 Confusion matrix

Confusion matrix is a classifier, which classifies the data into two classes, the positive class and negative class. The performance of a classifier has been evaluated using confusion matrix. The table 1 shows the data distribution using confusion matrix.

Table 1. Confusion matrix

True Positive(TP)	False Positive(FP)
False Negative(FN)	True Negative(TN)

3.3 Banana Leaf Disease Grading using ANFIS

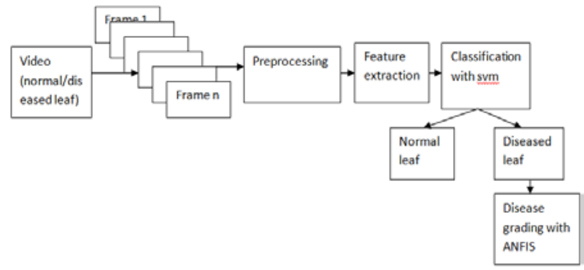

Fig.7. Disease grading system

In figure 7, SVM classifier will determine whether the leaf is diseased or not. If the leaf is diseased, then calculate the A_T (total area) and A_D (diseased area). Once A_T and A_D are known, the percent-infection (PI) is calculated by using ANFIS Toolbox [7].

3.3.1 Calculation of A_T and A_D

The resized image is converted to binary image such that the pixels corresponding to 'on' indicates the diseased area (A_D) and 'on' and 'off 'indicates the Total leaf area (A_T) .

3.3.2 Disease grading by Fuzzy Logic

Once A_T and A_D are known, the percent-infection (PI) is calculated by applying the formula (2).

$$PI= (A_D / A_T) *100 \quad (2)$$

The grade of the disease has to be determined from PI. Fuzzy Logic has been employed for this purpose. ANFIS system is used for the disease grading. Then grade the disease by using ANFIS Toolbox [7] along with a pre-determined grade score table. The Table 2 shows the grade table used for disease grading.

Table 2: Disease Scoring Scale for disease grading

Percent Infection	Disease Grade
0	0
Up to 1	1
1-10	2
10-20	3
20-40	4
40-100	5

3.5 Classification of Banana Leaf Diseases

The Banana Leaf diseases are classified using Multi layered Support Vector Machine. That is it detects whether the disease is Blacksigatoka disease or Panamawilt disease.

In Blacksigatoka, small yellow or brown colored lesions are present parallel to the vein of the leaf. Initially this lesions start from the vein.

In Panama wilt disease, the symptoms of this disease are yellowing ad withering of leaf also disease starts from border of the leaf. This leads to entire foliage within 2 or 3 days. It is leads to disease severity as high as 80-90%.

3.5.1 Multi layered Support Vector Machine

Multi layered Support Vector machine (MCS) classifies [4] the data into more than two classes. Here it takes three classes as shown in figure 8.

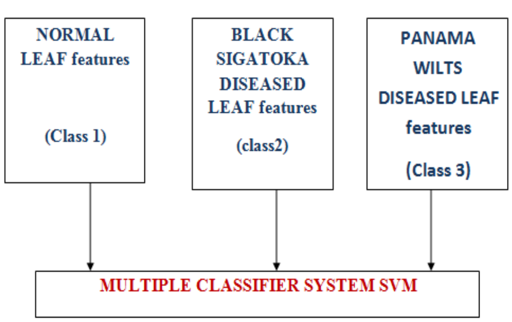

Fig.8. MCS system

4 RESULT ANALYSIS

For implementing the proposed system, here we used sample data from Mohandas College Banana farm by using Samsung Galaxy phone which contains50 video samples of both normal banana leaf and diseased banana leaf. Then image frames are generated using video reader function in mat lab 2013a. The image frames are selected randomly for testing and training as shown in table 3[9].

Table 3: video samples

Video of diseased leaf	Number of Frames/Video	Number of frames chosen
a	102	30
b	4719	30
Video of normal leaf	**Number of Frames/Video**	**Number of frames chosen**
d	216	10
e	2014	20
f	243	10
g	243	10

4.1 Disease detection using SVM & ANFIS

SVM or ANFIS is used to detect the disease. For that it accepts the input as still image or video.

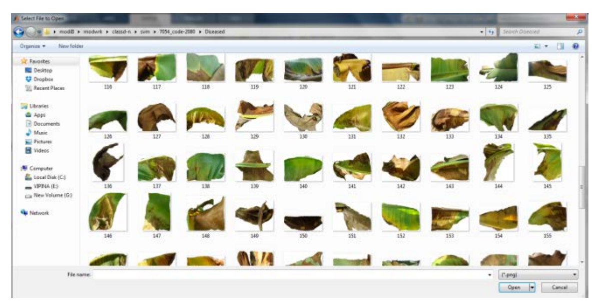

Fig. 9. Dataset for diseased leaf

From the dataset it selects the sample for disease detection. Figure 10 shows the classification result.

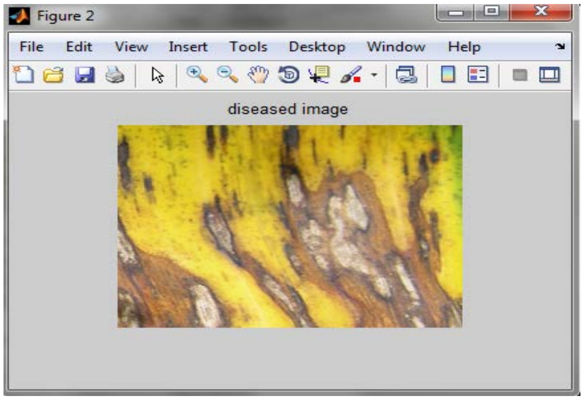

Fig. 10. Classification result for diseased leaf

Figure 11 shows the sample data (normal leaf) used for disease detection.

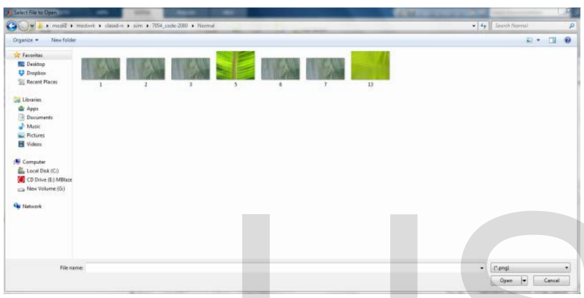

Fig. 11. Dataset for normal leaf

From sample data selects one image for feature extraction. Figure 12 shows the Testing result which shows the output is normal image.

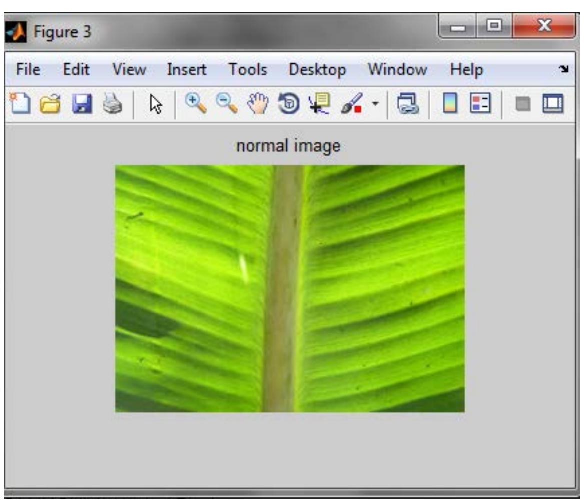

Fig. 12. Classification result for normal leaf

4.2 Banana leaf Disease Grading

Figure 13 indicate the input image need to be graded,

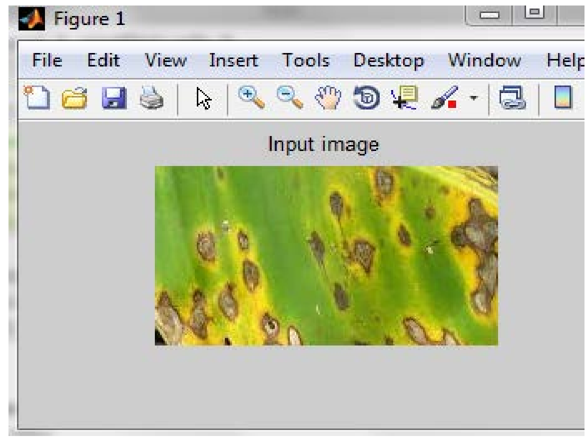

Fig.. 13. Diseased leaf

The Figure 14 shows the percentage of infection rate.

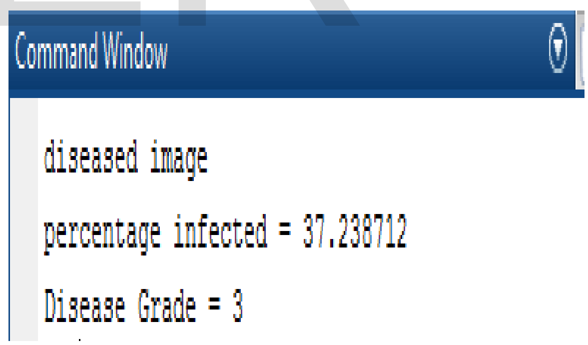

Fig. 14. Percentage of infection rate

4.3 Classifier comparison

4.3.1 Performance measurements of svm using confusion matrix

Table 4. Confusion matrix of svm

Type of Leaf	Normal	Diseased
Normal Leaf (60)	50	10
Diseased Leaf(70)	----	70

TPR= (correctly classified positives)/ (total positives)

=TP/ (TP+FN)

=50/ (50+0)

=1

FPR= (incorrectly classified negatives)/ (total negatives)

=FP/ (FP+TN)

=10/ (70+10)

=0.125

Specificity (Sp) =TN/ (TN+FP)

=70/ (70+10)

=0.875

Accuracy (Ac) = (TP+TN)/(TP+TN+FP+FN)

= (50+70)/ (50+70+10+0)

=0.923

i.e., Accuracy for SVM classifier is 92%.

4.3.2 Performance measurements of ANFIS using confusion matrix

Table 5. Confusion matrix of anfis

Type of Leaf	Normal	Diseased
Normal Leaf (60)	60	---
Diseased Leaf(70)	----	70

TPR= (correctly classified positives)/ (total positives)

=TP/ (TP+FN)

=60/ (60+0)

=1

FPR= (incorrectly classified negatives)/ (total negatives)

=FP/ (FP+TN)

=0/ (70+0)

=0

1-Specificity=0

Specificity (Sp) =TN/ (TN+FP)

=70/ (70+0)

=1

Accuracy (Ac) = (TP+TN)/ (TP+TN+FP+FN)

= (60+70)/ (60+70+0+0)

=1

i.e., Accuracy for ANFIS classifier is 100%.

4.4. Classification of Banana Leaf Diseases

Figure 15 shows the experimental result for Blacksigatoka disease.

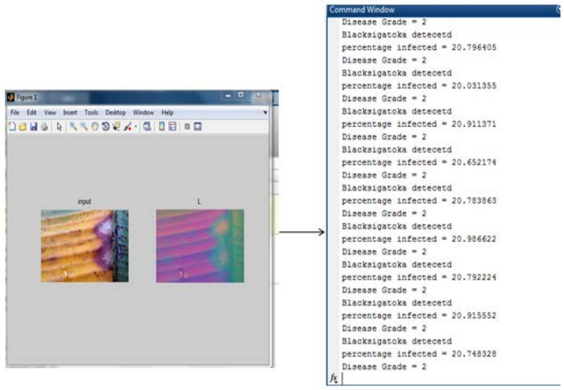

Fig. 15. Classification of Blacksigatoka disease

Figure 16 shows the experimental result for Panamawilt disease.

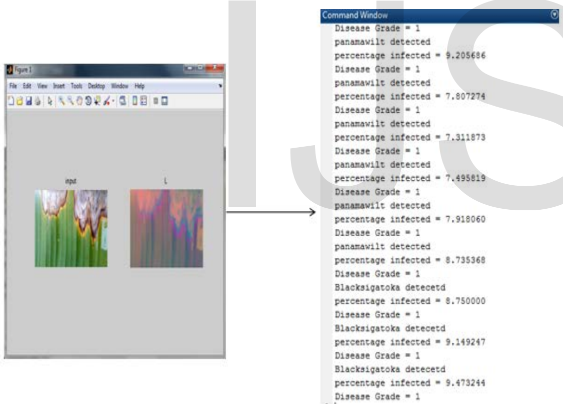

Fig. 16. Classification of Panamawilt disease

5 CONCLUSION

This work consists of identifying and grading the affected part banana Black Sigatoka disease and Panamawilt disease. It uses video as input. Banana leaf spot diseases were segmented efficiently according to color, texture and shape features. Initially Image segmentation is done, along with image analysis and important features are extracted and classification of diseases is performed using SVM classifier. If the leaf is diseased then disease grading is done using ANFIS classifier. Finally classifiers comparison has been performed with SVM or ANFIS classifiers using confusion matrix. The classifier comparison results 100 % accuracy for ANFIS and 92% for SVM. Which indicates ANFIS classifiers is the best classifier. The Multilevel SVM replaces Linear SVM which result 100 % accuracy for SVM. The observed results through experiments are found to be accurate and satisfactory.

REFERENCES

[1] Sanjeev S Sannakki, Vijay S Rajpurohit, Nargund, Arun Kumar R and Prema S Yallur Leaf Disease Grading by Machine Vision and Fuzzy Logic," IJCTA | SEPT-OCT 2011, Vol 2 (5.

[2] Sabah Bashir, Navdeep Sharma, "Remote Area Plant Disease Detection Using Image Processing IOSR Journal of Electronics and Communication Engineering (IOSRJECE), *ISSN : 2278-2834 Volume 2, Issue 6 (Sep-Oct 2012*.

[3]] P. Mohanaiah, P. Sathyanarayana L. GuruKumar, Image Texture Feature Extraction Using GLCM Approach, International Journal of Scientific and Research Publications, Vol.3, Issue 5, May 2013.

[4] N. Cristianini and J. Shawe-Taylor, "An introduction to support vector machines: And other kernel-based learning methods," Cambridge University Press, 2000.R.

[5] SVM classifier in matlab

[6] ANFIS Toolbox in matlab.

[7] K. Renugambal, B. Senthilraja (Asst Prof), "Application of Image Processing Techniques in Plant Disease Recognition", International Journal of Engineering Research & Technology (IJERT) IJERTV4IS030829, Vol. 4 Issue 03, March-2015.

[8] G.T.Hariharan,G.P.S.Hariharan,R.Vijay Anandh,"Crop Disease Identification using ImageProcessing",International journal of Latest Trends In Engineering and Technology(IJLTET), Vol.6, Issue 4, March 2016.

[9] Komal Bodkhe, Nisha Thakur, Shraddha Deshmukh, Asst. Prof. Prerana Jaipurkar," Analysis of Fungus in Plant Using Image Processing Techniques", International Journal of Recent Research in Mathematics Computer Science and Information Technology Vol. 2, Issue 1, pp: (12-18), Month: April 2015 – September 2015