Banana Plant Disease Diagnosis By Image Processing and Soft Computing Techniques

Mr.Athiraja A[1], Dr.Vijayakumar P[2]
Assistant Professor[1], Professor[2]
Rathinam Technical Campus[1], Karpagam College of Engineering[2]
Coimbatore, Tamilnadu, India.

Abstract:

As it is observed that the banana production is affected by various disease condition and causing huge loss to the poor farmers. By using modern technology of image process and soft computing, these can be identified at the earlier stage and appropriate precautions can be taken to avoid further damage and hence increase in healthy production. Through the pre-processing technique, image is input to get normalization and median filter is done to remove the noise. Then segmentation process is done for feature extraction, followed by classification techniques. In this classification techniques two algorithms are used, that is the perceptron and case based reasoning for diagnosis of disease. Then fuzzy logic is used for making decision. Then system evaluation Receiver Operating Characteristics (ROC) curve is used. Analysis shows Case based reasoning (CBR) algorithm is better than perceptron algorithm.

Key words: segmentation, perceptron algorithm, CBR, ROC, fuzzy logic.

1.INTRODUCTION:

Banana serves as a staple as well as a cash crop in India. The Banana cultivation has been contributing a lot in the rise of economy[2]. The biggest advantage of Banana is it's availability throughout the year and its production is high in Asian countries.

There are different varieties of Banana available and it's based on the region of cultivation. From a study done by BAPNet [1&3], there are nearly 16 varieties of **Musa Balbisiana** available from Indian cultivation grounds. Since there are a lot of varieties, we also experience a lot of different diseases in these plants [4].

The common disease is the leaf spot disease. The banana plants from Kodaikanal and Anaimallai hills prove with zero leaf-spot disease type[5]. Based on the region, soil, and humidity conditions the plant experiences different diseases[6].

From a recent survey it is found that the second largest banana production in India is from Tamil Nadu.
Banana plant disease spreads very fast and thus damaging the crop, this leads to crop destruction, low production, and a fall in economy.

Research Solution: In this research work, the crop diseases are being identified by image processing. This saves time and suggests the best method to adopt to overcome the disease at its earliest stage. In this project the ROC curve is used to express the whole process of evaluation.

2.DATA COLLECTION

The data is collected from Sivangangai, Madurai, Virudunagar, Thani, Dindugal, Coimbatore, Tripur, Salem, Namakkal, and Thirunelvi.

The following diseases are taken for study and observation:

- Panama wilt
- Leaf spot
- Anthracnose
- Cigar-end tip rot

- Crown rot
- Virus disease

Figure (1a)

Figure (1b)

Random sample of data from around 40 farmers have been taken. Sample images are shown in **Figure (1a) and Figure (1b)**. The very common varieties of banana available in Tamil Nadu are Rasthali, Poovan, Red banana, Karpporavalli, Matti, Ney Poovan, Vayal Vazhai, Nadu, Ottu.

3.METHODLOGY

The following **Figure (2)** block diagram indicates the methodology of the proposed system used to identify the banana plant disease by image processing and soft computing techniques. By this technique, the disease can be prevented at the earliest stage.

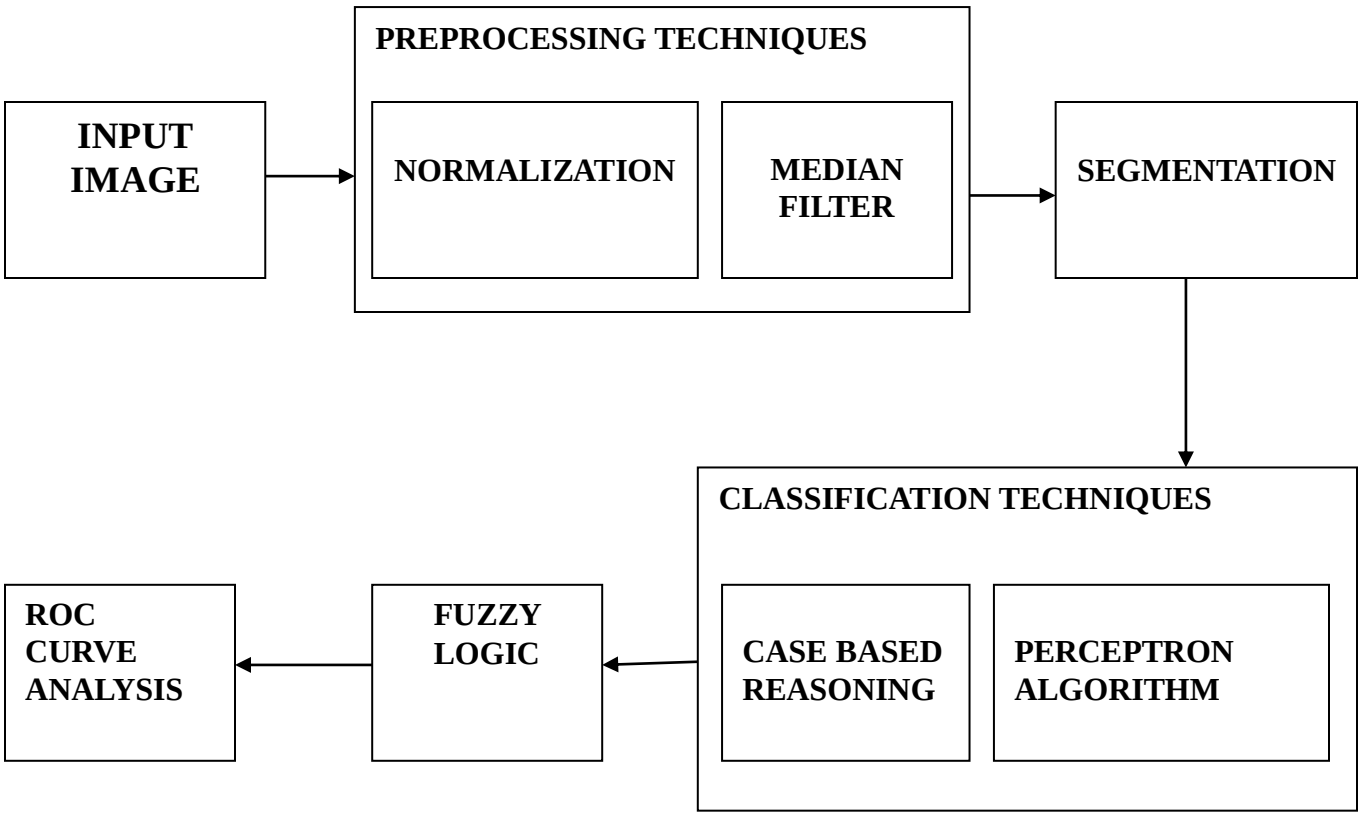


Figure (2) Block diagram

4. INPUT IMAGE

The input images are collected from Research centres, Agriculture Universities and directly from the agriculture farms. The input images are fed in JPG format normally in windows. It is stored in matrix format. Input image for various diseases as shown in **Table (1).**

PANAMA WILT 	**LEAF SPOT** 	**CROWN ROT** 
ANTHRACNOSE 	**CIGAR-END TIP ROT** 	**VIRUS DISEASE** 

Table (1) input image

5. PRE-PROCESSING:

Pre-processing is the process carried out to eliminate the noise from the captured image.
Pre-processing method consists of two steps:

- Normalization
- Median filter

Normalization:

The original image will have its own colour and these colours have their own pixel range. These images have both colour and intensity information. By normalization, the original coloured image is converted into a grey image. This grey image has intensity information only. RGB pixel ranges for both the images are calculated using equation (1) and they are used for comparison for the results.

I_N = (I - Min) ((newMax-newMin) / (Max - Min)) + newMin Equation (1)

Median Filter:

Median filtering is a technique mainly used to reduce noise from the image. It is a little different from usual pre-processing technique. The noise in the boundary line is undisturbed and it focuses more on the centre pixel's noise.
It is also one of the smoothing techniques. Smoothening technique is by which the noise is being reduced to smooth patches and smooth regions. Pre-processing output image for various diseases are shown in **Table (2).**

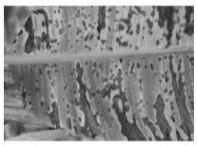
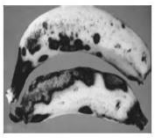
 **PANAMA WILT**	 **LEAF SPOT**	 **CROWN ROT**
 **ANTHRACNOSE**	 **CIGAR END TIP ROT**	 **VIRUS DISEASE**

Table (2) Pre processing output

6. FEATURE EXTRACTION:

The main things that are got out of these feature extraction are effectiveness and efficiency of analysed data. There are three feature extraction techniques commonly used:

- Shape extraction
- Colour extraction
- Texture extraction.

In this research, the colour extraction technique is used.

COLOUR EXTRACTION:

The colour extraction method uses three methods of extraction

- RGB
- HSV
- YCbCr

In colour extraction method, each pixel is having its individual values. The values of RGB from the pixels of the image are used as the colour extraction inputs. While extracting the colours from the image the values of 0's (zeros) and 1's (ones) play an important role. The 0's and 1's gives the information as to which part is to be extracted. The part with 1's as the output is the one which is the essential and it has to be extracted. The output of the colour feature extraction method is shown in **Figure (3).**

Figure (3) feature extraction output

7. IMAGE SEGMENTATION:

In image segmentation the colour feature is taken into account. It is known that the colour of a healthy leaf differs from that of a diseased leaf. So there is a difference in the colour of the leaf and the varied pixel values of that colour are taken as the readings to perform the image segmentation.

$$\text{Seg}(x,y) = \begin{cases} 1 \text{ if } I(x,y) > T \\ 0 \text{ if } I(x,y) < T \end{cases} \quad \text{Equation (2)}$$

If the threshold value is greater than the pixel value then the value given to it is 1.
If the threshold value is smaller than the pixel value then the value given to it is 0.
Here the image is segmented into two using equation 2 and processed output image for various diseases as shown **Table (3)**. Threshold value for various diseases is listed below

- Threshold values of Panama wilt diseased leaf is 160 to 210
- Threshold values of Leaf spot is 60 to 100
- Threshold values of Anthracnose is 30 to 90
- Threshold values of Cigar-end tip rot is less than 70
- Threshold values of Crown rot is less than 60
- Threshold values of Virus disease is less than 85

 **PANAMA WILT**	 **LEAF SPOT**	 **CROWN ROT**
 **ANTHRACNOSE**	 **CIGAR END TIP ROT**	 **VIRUS DISEASES**

Table (3) Segmentation output

8. PRECEPTRON ALGORITHM:

The perceptron algorithm is a supervised machine learning algorithm. It is a type of linear classifier and it gives the possible output. In simple, the perceptron algorithm is used to predict the possible output. It performs the prediction based on the set of weights of the feature vector.

The Perceptron is a classifier which maps its input x(positive value vector) to an output value y (either 0 0r 1) calculated using following equation(3):

$$y = \begin{cases} 1 & if \;\; w.x + b > 0 \\ 0 & otherwise \end{cases} \quad \text{Equation (3)}$$

LEARNING ALGORITHM

1. Initialize weights and threshold. Note that weights may be initialized by setting each weight node $w_i(0)$ to a small random value or to 0. In the example below, we choose the former.

2. For each sample j in our training set T, perform the following steps over the input xi and desired output yi:

2a. To calculate the actual output using equation (4): $yi(t) = f[w(t).xi]$ Equation (4)

$$yi(t) = f[w0(t) + w1(t)xi,1 + w2(t)xi,2 + \cdots + wn(t)xi,n$$

2b. Adapt weights calculated using equation (5): $wi(t+1) = wi(t) + \alpha(di - yi(t))xi,j$

Equation (5) for all nodes 0<i<n

Step 2 is repeated until the iteration error is less than a error threshold value set by user, or a number of iterations have been completed defined by user.

Execution of the Perceptron algorithm following steps

Step1: Pre-processed mammogram image is converted into the two column vector (x1, x2).

Step 2: Two column values is converted into binary value based on the following assumption. If (x1 and x2 > 180) then assign to 1 otherwise -1.1-Represents the Abnormal cells and -1- Represents the Normal absent cells.

Step 3: To find out the target value as shown in the following **Table (4)**

X1	X2	T
1	1	1
1	-1	1
-1	1	1
-1	-1	-1

Table 4 OR logic function

Here using the OR logic function to find out the target value.

Step 4: To calculate the net output value using the equation (4)

Step 5: The step 4 is continued up to termination condition. These conditions are

- Number of epoch
- Actual value always equal to target value.
- There is no weight change.

Step 6: Finally classify the Normal cells and abnormal absent cells as shown in **Table (5).**

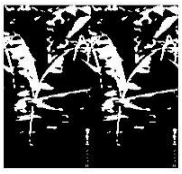
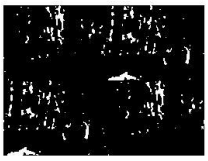
 **PANAMA WILT**	 **LEAF SPOT**	 **CROWN ROT**
 **ANTHRACNOSE**	 **CIGAR-END TIP ROT**	 **VIRUS DISEASE**

Table (5) Perceptron algorithm output

9. CASE BASED REASONING:

Case based reasoning is also a type of prediction technique. This is different from Perceptron algorithm. Here the data inputs are taken from the existing database available. The data from the previous year databases are used

as the data inputs to do the prediction. Case based reasoning is an unsupervised machine learning algorithm. The case based reasoning is classified into two.

Normal and abnormal are the two basic classifications. The abnormal pixel value is further classified into six such as initial, very small, small, medium, high, very high. Since the threshold values are set and taken from the existing database the output values are expected to be very accurate as shown **Table (6).**

Based on the classification, if the result is in initial stage then it is corrected immediately. In-case of medium or high sections then serious actions is taken whether to remove the plant from that region or to treat the plant with fertilizers. Output of CBR for various diseases are shown in **Table (7).**

Classification	**Panama Wilt**	**Leaf Spot**	**Crown Rot**	**Anthracnose**	**Cigar-end Tip Rot**	**Virus Disease**
Normal	<160 & >210	<60 & >100	>70	<30 & >90	>60	>85
Abnormal	160-210	60-100	<70	30-90	<60	<85
1.initial	160-165	60-65	0-9	30-39	0-9	0-15
2.very small	166-170	66-70	10-19	40-49	10-19	16-30
3.small	171-175	71-79	20-29	50-59	20-29	31-40
4.medium	176-180	80-84	30-49	60-69	30-39	41-59
5.high	181-194	85-89	50-59	70-79	40-49	60-69
6.very high	195-210	90-100	60-69	80-89	50-59	70-84

Table (6) Threshold value for Various Diseases

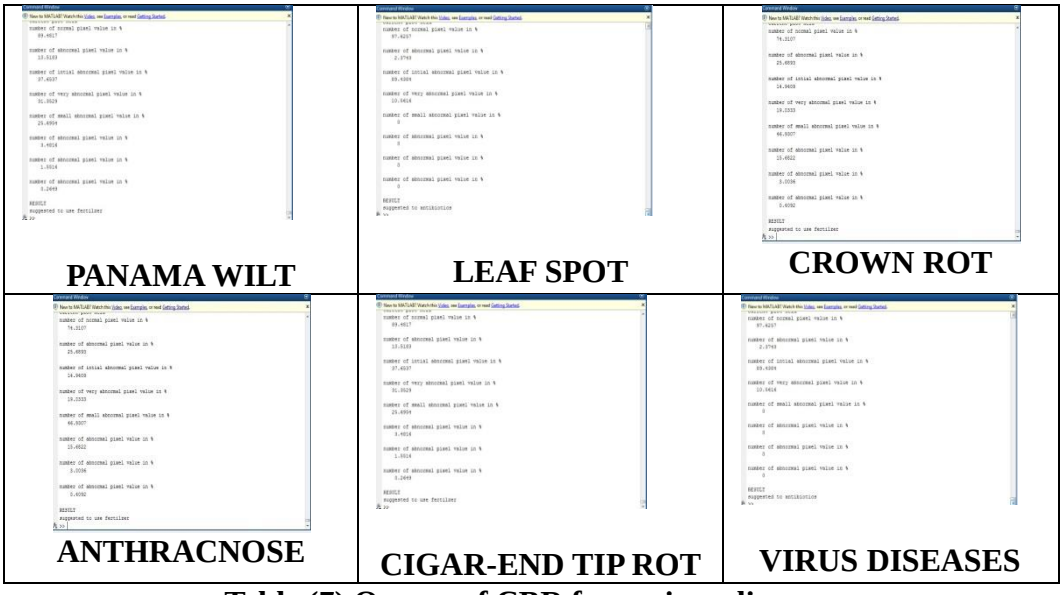


Table (7) Output of CBR for various diseases

10. FUZZY LOGIC:

Fuzzy logic is a many-valued logic technique, where the results are given in degrees of truth. The degrees of truth are between 0-1. In crisp

logic, the results are either 1 or 0. Unlike crisp logic, fuzzy logic gives the results as degrees of truth from 0, 0.25, 0.5 and 1. For abnormal cells, if-then rule is used. Output image of fuzzy logic are shown in F**igure (4)**

- If the abnormal value > 60 then it is suggested to remove the plant from that region:
- If the abnormal value >20-40 then it is suggested to treat the plant with fertilizers.

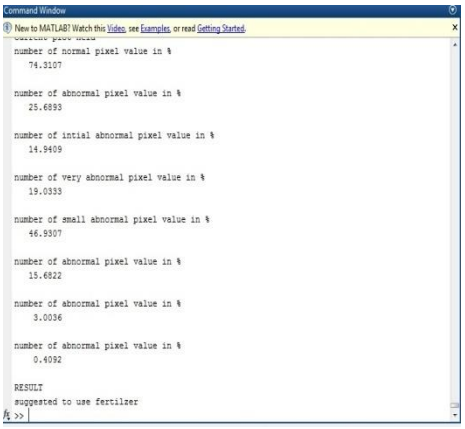


Figure (4) Fuzzy Logic Output

11. ROC CURVE:

The ROC curve is the extension of receiver operating characteristics. This ROC curve is a graphical representation which plots the values obtained into true positives out of positives and false positives out of negatives as shown **Figure (5).**

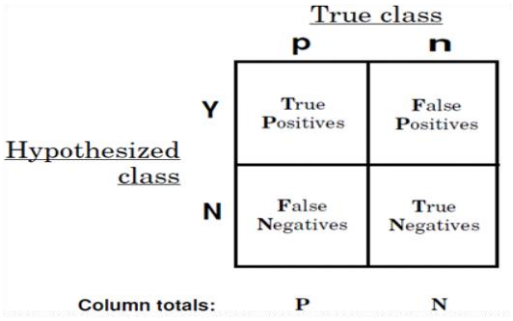


Figure (5) ROC true positive and false negative

Sensitivity or true positive rate (TPR) calculated using following equation (6)

$$TPR = \frac{TP}{P} = \frac{TP}{(TP+TN)}$$ Equation (6)

Specificity (SPC) or True Negative Rate calculated using following equation (7)

$$SPC = \frac{TN}{N} = \frac{TN}{FP+TN}$$ Equation (7)

To plot the roc curve sensitivity vs. specificity is shown in below figure (6) this for out of the panama wilt disease.

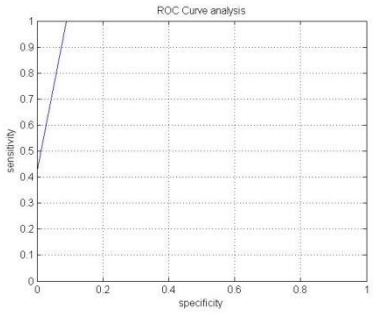


Figure (6)

12. RESULT ANALYSIS:

To measure the each method accuracy (ACC) using the following equation (8)

$$ACC = \frac{TP+TN}{P+N}$$ Equation (8)

The accuracy value of the proposed method is shown in following **Table (8)**

Diseases	Image segmentation	ANN	CBR
Panama wilt	82	85	92
Leaf spot	86	87	93
Anthracnose	83	80	91
Cigar-end tip rot	86	81	94
Crown rot	85	84	92
Virus disease	74	80	90

Table (8) Accuracy Value For Proposed Methods

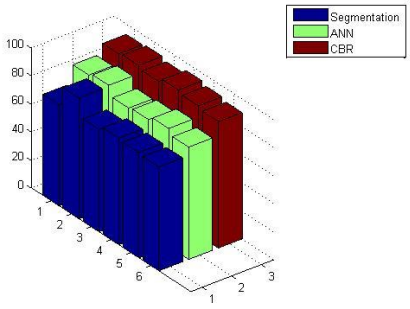


Figure (7)

The above table, value and graph as shown in **Figure (7)** indicate the accuracy of proposed methods. Among these three methods CBR provides better accuracy.

12. CONCLUSION:

By using image processing, the diseased parts are identified earlier and better treatment measures are given. Rather than the whole plant dyeing, by these modern techniques, it could actually prevent the whole plant from perishing and help to regain the plants natural self. In this research work, three methods are used. Image segmentation method accuracy average value is 82.3%, Perceptron algorithm accuracy average value is 82.6% and CBR accuracy value is 92%. Case Based Reasoning method provides better accuracy compared to image segmentation method and Perceptron algorithm.

13. REFERENCE:

[1]. A.B. Molina, V.N. Roa and M.A.G. Maghuyop, Advancing banana and Advancing banana and plantain R & D in Asia plantain R & D in Asia and the Pacific - V and the Pacific - Vol. 10 Proceedings of the 10th INIBAP-ASPNET Regional Advisory Committee meeting held at Bangkok, Thailand -- 10-11 November 2000.

[2]. A.B Molina, BP Joven, VGO Sinohin, JD Generoso , Advancing Banana And Plantain R&D In Asia And The Pacific Proceedings of the 7th BAPNET Steering Committee Meeting Hanoi, Vietnam 2-5 November 2010.

[3]. A.B. Molina, L.B. Xu, V.N. Roa, I. Van den Bergh and K.H. Borromeo, Advancing banana and plantain R&D in Asia and the Pacific - Proceedings of the 3rd BAPNET Steering Committee meeting held in Guangzhou, Vol. 13, China, 23-26 November 2004.

[4]. A.B. Molina, V.N. Roa , Advancing Banana And Plantain R&D In Asia And The Pacific Proceedings of the 6th BAPNET Steering Committee meeting in Tiruchirapalli, Tamil Nadu, India 22-25 October 2008.

[5]. Juliane Henderson (UQ/DPI&F), Kathy Grice (DPI&F), Marie-Françoise Zapater (CIRAD), Françoise Careel (CIRAD), Mycosphaerella Spp. Causing Leafspot Disease On Bananas Advanced Methods Supplement, October 2008.

[6]. Identification and Integrated PEST Management in Banana and Plantain Magdalena and Uraba Colombia June 2009 ISBN 978-958-99167-0-4