CLASSIFICATION OF PADDY LEAF DISEASES USING SHAPE AND COLOR FEATURES

Suman T[1], Dhruvakumar T[2]

[1]*M.Tech student, Signal processing, Siddaganga Institute of technology, Tumakuru, Karnataka, (India)*
[2]*Asst. Professor, Dept. of Electronics and Communication, Siddaganga Institute of Technology, Tumakuru, Karnataka, (India)*

ABSTRACT

Farmers experience great difficulties in early disease detection and it is a major challenge in agriculture field. Automation of essential processes in agriculture is becoming widespread, especially when fast action is required. In this work suitable preprocessing techniques are applied and using histogram plot, normal and diseased leavesare classified. For the diseased leaf, shape and color features are extracted and the combined features of color and shape are used for the classification of bacterial leaf blight, brown spot, narrow brown spot and rice blast diseases using SVM classifier.

Keywords*: *SVM classifier, shape features, color features, bacterial leaf blight, brown spot, narrow brown spot and rice blast.

I INRODUCTION

Agriculture has always been the mainstay in economy of most of the developing countries, especially the ones located in South-Asia. The purpose of agriculture is not only to feed ever growing population but it's a solution to solve the problem of global warming and it is an important source of energy. India is an agricultural country wherein most of the population depends on agriculture. Research in agriculture is aimed towards increase of productivity and quality of the food at reduced expenditure, with increased profit. Therefore, detection and classification of diseases is an important and urgent task.

The amount of crops that are damaged every year, due to adverse climatic conditions or invasion of pathogens, can never be neglected. Hence, it is important for the farmers to detect the growth of disease in plant at an early stage, and take necessary steps in order to prevent it from spreading to others parts of the field.

Rice is a globalized staple food. It is one of the three leading food crops in the world which makes it a more significant food item worldwide [1]. Rice (Oryzae Sativa) is considered as the main crop in the east India and believed to be the second central crop after wheat, in the world. In a third world country like India where the major staple food is "Rice" where life of many people, economy of the country is related to the production of paddy. Any negative effect on the yield in unwanted. The paddy production can be hampered as effect of some mechanical damage, nutritional deficiency, genetically disorder, climatic conditions etc [5]. But the major problem is disease causing by macrobes [2] and microbes. The disease is easily recognized by their symptoms-changes of the plants.

The efforts to increase the quantity and quality of rice production to satisfy the increasing needs of rice in India experienced several obstacles, one of which is the attack of the diseases on paddy fields. To control these diseases and to minimize the impacts of the attacks, the diseases must be identified quickly. Computer vision is a potential solution to tackles this problem. One way to identify the diseases in plants is by observing the physical changes (diseases spots or lesions) caused by chemical changes in the sick plants. The images of these spots can be processed and used to recognize the diseases quickly, easily, and inexpensively.

The most two common diseases in the North East India are named as Leaf Blast and Brown Spot. The samples of the infected rice leaves have been collected from different parts using Nikon COOLPIX P4 digital camera. Acquired images transformed to Hue Intensity Saturation (HIS) model for segmentation [3]. Entropy based bi-level thresholding method has been invoked for segmenting the images to facilitate identifying the infected parts of the leaves [4].

The RGB color images of paddy leaf are captured using a Canon Power Shot G2 digital camera [5]. The image segmentation based on gray-level threshold segmentation is adapted and the binary image is gained. The main objective of segmentation process is to obtain the binary image with less noise or noise free [6]. The RGB image is converted into a binary image using threshold method. Local entropy threshold method of Eliza and Chang [7] and Otsu method is used for the segmentation [8]. An occurrence matrix is generating from the input image in accordance with probability distribution needed for entropy measures. Five characteristics of lesion i.e., percentage, lesion type, boundary color, spot color, and broken paddy leaf color were tested for the classification task Color is an important sign in recognizing different classes [9].

Four characteristics of lesion type, boundary color, spot color, and broken paddy leaf color were tested for used to establish the classification system. The ratio of height and width of the lesion spot provided a unique shape characteristic for determining the type of the lesion [6]. Generally, the color difference is evaluated using the distance between two color points in a color space. The most common distance is Euclidean distance [6]. Our proposed technique is based on the CIELab color space, which is a uniform chromaticity color space to get boundary color, spot color and broken leaf color. It is known that Euclidean distance of two colors is proportional to the difference that human visual system perceived in the CIELab color space [7][10].

A probabilistic neural network (PNN) is nonparametric classifiers [11]. PNN work faster than the back propagation neural network, even up to 200,000 times faster as it only needs one iteration of training process. Training and testing data were splitted using 5-fold cross validation. The results were presented using confusion matrix for further analysis.

In this present work four different types of paddy diseases are considered namely, brown spot, narrow brown spot, bacterial leaf blight and leaf blast. Here, bacterial leaf blight and narrow brown spot has the similar shape and rice blast, brown are almost similar in color, hence both the color and shape features are used for the classification purpose.

II DISEASES OF PADDY LEAVES

Rice plant is distress from many diseases [1], [2] the main diseases are caused by bacteria and fungus. The RGB normal rice leaf is shown in Fig 1(a). The diseases which are considered in this work are listed below

2.1 Brown spot

It is caused by the virus named as Cochliobolusmiyabeanus &Helminthosporium. Brown spot diseased leaf is shown in Fig 1(b). The main symptoms are

- Initially appears as brownish spots on leaves.
- Later, it becomes Oval shaped foliar spots with yellow halo.
- The spots are brown, with greyish centers when fully developed.
- Appear in leaf blade & sheath

2.2 Narrow brown spot

It is caused by the virus named as Cochliobolusmiyabeanus&Helminthosporium. Narrow Brown spot diseased leaf is shown in Fig 1(c). The symptoms are similar to the brown spot but these are long and narrower when compared to brown spot diseased part.

2.3 Bacterial leaf blight

It is caused by the bacteria named as Xanthomonasoryzae_andpv. Oryzae. : Bacterial leaf blight diseased leaf is shown in Fig 1(d). The main symptoms are

- Water soaked lesions move from tip downwards on the edges of leaves.
- Gradually symptoms turn into yellow and straw colored stripes with wavy margins.
- In early morning in humid areas yellowish, opaque, turbid drops of bacterial ooze may be seen.

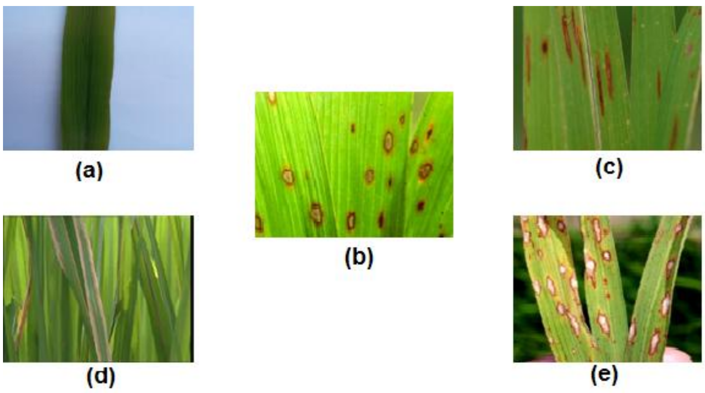

Figure 1: RGB image of normal leaf (a), (b)-(e) diseased leaves of brown spot, narrow brown spot, bacterial leaf blight,, rice blast respectively.

2.4. Rice Blast

It is caused by the fungus named as Pyriculariagrisea. Rice blast diseased leaf is shown in Fig 1(e). The main symptoms are

- Start as small water soaked bluish green specks.
- Leaf spots are typically elliptical (football shaped), with gray-white centers and brown to red-brown margins. Fully developed leaf lesions are approximately 0.4 to 0.7 inch long and 0.1 to 0.2 inch wide.

III PROCESS DESCRIPTION

Different types of diseased and normal paddy leaves are captured using digital camera. After that an image is cropped manually in such a way that it contains both healthy and diseased part in it. It is the only manual work involved in this process and this image is given as input to the further process. The block diagram of proposed work is shown in Fig 2.In image pre-processing green plane is extracted to enhance the diseased part. The major objective of preprocessing stage can be to reduce the amount of noise present in the document and to reduce the amount of data to be retained. Median filter is used in order to remove the noise present in the acquired image. Diseased part of the leaf image need to be extracted by using suitable segmentation technique. The features of the segmented image are extracted in order to recognize the different diseases. Then the data extracted in the features are stored in the database.

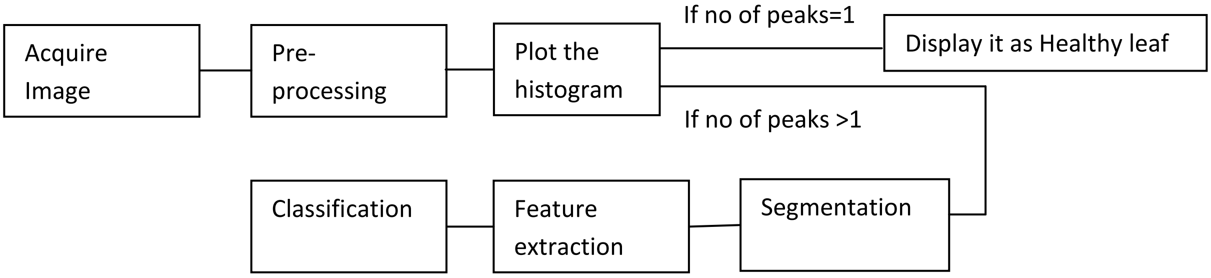

Figure2: Block diagram of proposed leaf disease detection and classification

3.1 Preprocessing

Image preprocessing is the name for operations on images where it can be used to improve the image data that removes the background noise and also suppress the undesired distortion. Through various image preprocessing steps, image features for processing and analysis are enhanced. The preprocessing steps involved in this work are represented in Fig.3.

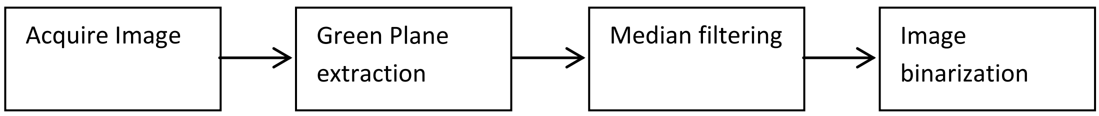

Figure 3: Block diagram of preprocessing steps involved in the proposed work

3.1.1. Acquisition of images

Acquire the images of normal and diseased leaf using digital camera with a white background. Images are taken with a white background in order to avoid the reflections while capturing the images.

3.1.2. Green plane extraction

Acquired RGB imageis used for further preprocessing involved in the proposed methodology. It can be easily observed that greenness of the leaf is more affected when the infection is occurred in leaf. Extract the green color component in order to enhance the affected portion of the leaf. Instead of considering only the green values, intensity of the original gray scale image is subtracted by the green value so that the spot detection is invariant of the brightness and age of the leaves.

3.1.3. Median filtering

Median filter is nonlinear in nature and it is a smoothening filter. Here, it is used to preserve the edge information as it had a less blurring effect when compared to the other smoothening filters after applying on to the image. Median filter is applied to remove unnecessary spots. As a result, a noise free grayscale image is produced. It replaces the value of the center pixel, by the median of the gray levels in the image area enclosed by the filter. In order to perform median filtering, first window is moved and all the pixels enclosed by the window are sorted. After then median is computed and this value is assigned to center pixel. If the number of elements in 5X5 window is odd, middle value is assigned as median value, else average value is assigned as median value, and else average of two middle values is assigned as median value. Median filtering operation is shown in Fig.4.

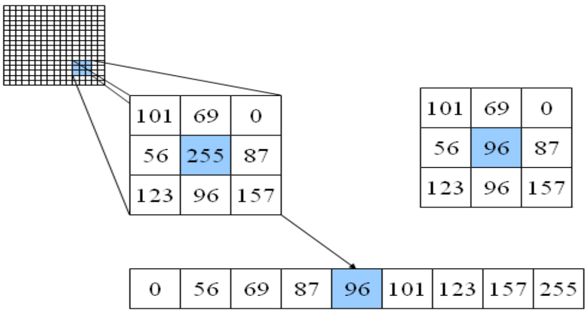

Figure 4: Kernel showing the median filtering operation

3.1.4. Binarization

Thresholding is used to convert a gray scale image into the binary image. Thresholding technique replaces all pixels of an input image with black pixels if the intensity $I_{x,y}$ less than some fixed threshold T or a white pixel if the intensity is greater than the fixed threshold T.So that the resultant image is binarized by applying the thresholding.

3.2. Classification of healthy and diseased leaf

For an uninfected leaf the distribution of color is nearly uniform while for the diseased leaf the distribution of the color is not uniform because the pixel values of the diseased leaves varies widely and differs from the normal intensity value of pixels. In order to distinguish between the uninfected and diseased leaves a histogram approach has been used. Histogram represents the probability of occurrences of different gray levels in the image. The high probability of a particular gray level corresponds to a peak in the histogram.

Binary image is multiplied with the original gray scale image and the histogram is plot to the resultant image. Count the number of pixels lie in the intensity value greater than zero. If the sum is obtained then the result is unhealthy leaf. If the sum is zero then it is healthy leaf.

In the case of healthy leaf intensity value is lie in only one value so the peak is one and the result will be displayed as healthy leaf. In the case of an diseased leaf pixel intensity spread in all intensities so the peak is greater than one and hence it is displayed as unhealthy leaf.

3.3. Segmentation

Image segmentation is the process of partitioning a digital image into multiple segments. The goal of segmentation is to simplify and change the representation of an image into something i.e. more meaningful and easier to analyze. The result of image segmentation is a set of a segment i.e. collectively cover the entire the image or set of contour extracted from the image. The main idea of the image segmentation is to group pixels in homogeneous regions and the usual approach to do this is by common feature. Image segmentation is the process of dividing the given image into regions homogenous with respect to certain features for image segmentation. 8-connected component analysis is used for the segmentation of the diseased leaf.

IV FEATURE EXTRACTION

The main aim of feature extraction is to extract the information that can be used to determine the meaning of the given sample. Shape, color and texture features are the main features that are included in image processing. Here, shape of the infected region and color changes in the lesion area are considered as the features for the classification different kinds of diseases. The color features are influenced by outside light and different diseases of plant leaves had same shape features, hence both color and shape features are considered as the characteristic values of classification of different types of diseases.

4.1. Shape features

Principal Component Analysis is used for shape feature extraction. PCA transforms correlated variables into uncorrelated variables retaining maximum amount of variations. This helps to operate on data and make predictions. Each image is converted into vector and stored as columns of matrix $P \times N$. Mathematically principal components are obtained by subtracting mean from each column. The resultant matrix is given by B.

$$B = [\hat{X}_1, \hat{X}_2 \dots, \hat{X}_N] \qquad (1)$$

Then the covariance matrix S is obtained by

$$S = \frac{1}{N-1} B^T B \quad (2)$$

Thus the dimension of the feature vector reduces to $P \times P$. The eigen values and eigen vectors are calculated using $P^T SP = D$. Where D is the diagonal matrix with eigen values. And P is is a matrix of eigen vectors. Eigen vectors are arranged according to the descending values of the eigen values. The weight matrix W is determined as $W = PB^T$and this weight matrix is used as features for classification.

4.2. Color features

In image processing color features plays very important role and an important sign in recognizing different classes. These color features are very helpful when investigating the lesion for early diagnosis. Here, "Grid based color moments" are used as a feature vector. Compute the color features for a given image using following steps

- RGB image converted into HSV color spaces
- An image is uniformly subdivided into 3X3 blocks
- Compute mean color (H/S/V) for each of the nine blocks

$$x' = \frac{1}{N} \sum_{i=1}^{N} x_i \quad (3)$$

Where N is the total number of pixels within each block,x_i is the pixel intensity in H/S/V channels.

Mean is considered as one of the feature as it measures the average intensity value.

- Compute variance for each block and for each channel

$$\sigma^2 = \frac{1}{N} \sum_{i=1}^{N} (x_i - x')^2 \quad (4)$$

Variance has the capability of measuring the variability as the intensity level get changes at the edges of the images by large value variance can be used to sharpen the edges.

- Compute the skewness for each block of(H/S/V)

$$\sigma^2 = \frac{\frac{1}{N} \sum_{i=1}^{N} (x_i - x')^3}{(\frac{1}{N} \sum_{i=1}^{N} (x_i - x')^2)^{\frac{3}{2}}} (5)$$

Usually skewness is used is used for judging the image surface. It can be positive, negative or zero. Skewness measures the asymmetry of the image.

Each block will have 3+3+3=9 features, and thus the entire image will have 9x9=81 features. Before we use SVM to train the classifier, we first need to combine the shape and color features for claasification.

V SVM CLASSIFIER

SVM is a powerful discriminative binary classifier which models the decision boundary between two classes as a separating hyper plane. This hyper plane tries to split, one class consists of the target training vector (labeled as +1), and the other class consists of the training vectors from an impostor (background) population (labeled as −1). Using the labeled training vectors, SVM optimizer finds a separating hyper plane that maximizes the margin of separation between these two classes and it is shown in Fig.5.

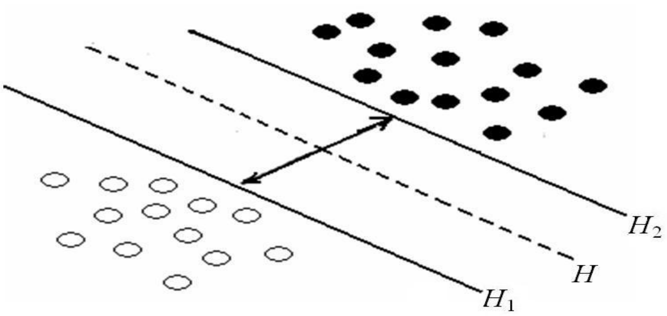

Figure 5: The optimal plane of SVM in linearly separable condition

Formally, the discriminative function of SVM is given by

$$f(x) = \sum_{i=1}^{N} \alpha_i t_i K(x, x_i) + d \qquad (6)$$

Here $t_i \in \{+1, -1\}$ are the ideal output values, $\sum_{i=1}^{N} \alpha_i t_i = 0$ and $\alpha_i > 0$. The support vectors x_i, their corresponding weights α_i and their basis term d, are determined from the training set using an optimization process. The kernel function $K(.,.)$ is designed so that it can be expressed as $K(x, y) = \varphi(x)^T \varphi(y)$, where $\varphi(x)$, is a mapping from the input space to kernel feature space of high dimensionality. The kernel function allows computing inner products of two vectors in the kernel feature space. In a high-dimensional space, the two classes are easier to separate with a hyperplane. Intuitively, linear hyperplane in the high-dimensional kernel feature space corresponds to a nonlinear decision boundary in the original input space. The most widely used kernel functions are: Linear kernel, the Radial Basis Function function kernel (RBF kernel), the sigmoid kernel. Since the real world problem deals with multi class classification. This problem can be solved using two approaches: One-Against-All (OAA) approach and One-Against-One (OAO) approach.

VI EXPERIMENTAL RESULTS

For the detection and classification of plant leaf diseases, acquire the images of normal and diseased leaf using digital camera with a white background, and this acquired image is passed through different preprocessing steps. Extraction of green color component of four different diseases i.e., bacterial leaf blight, brown spot, narrow brown spot and rice blast respectively is shown in Fig.6. Median filter is applied to remove unwanted spots and the results when applies to four diseases are shown in Fig.7. Intensity enhanced image in order to enhance the diseased part is shown in Fig.8. Fig. 9 shows the results of binarized image after applying the thresholding. Classification of healthy and disease leaf for healthy, bacterial leaf blight, brown spot, narrow brown spot and rice blast respectively are shown in Fig. 10. Results of segmentation using 8-connected component analysis aregiven in Fig.11.

Figure 6:Green plane extraction of bacterial leaf blight, brown spot, narrow brown spot and rice blast respectively.

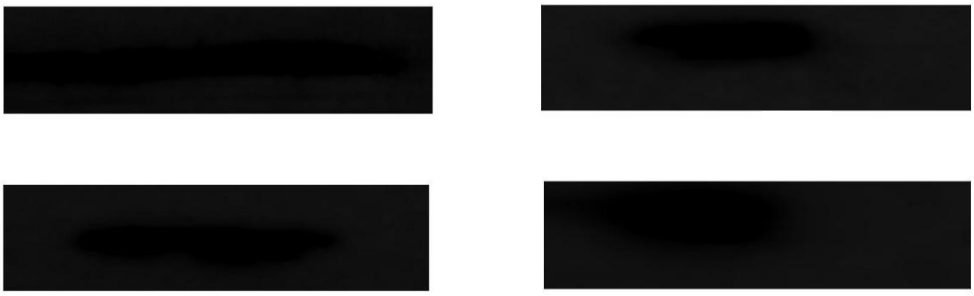

Figure 7:Median filtered results of bacterial leaf blight, brown spot, narrow brown spot and rice blast respectively.

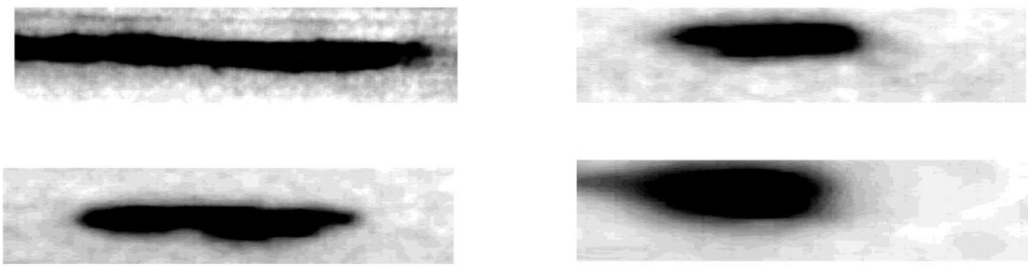

Figure 8:Intensity enhanced results of bacterial leaf blight, brown spot, narrow brown spot and rice blast respectively.

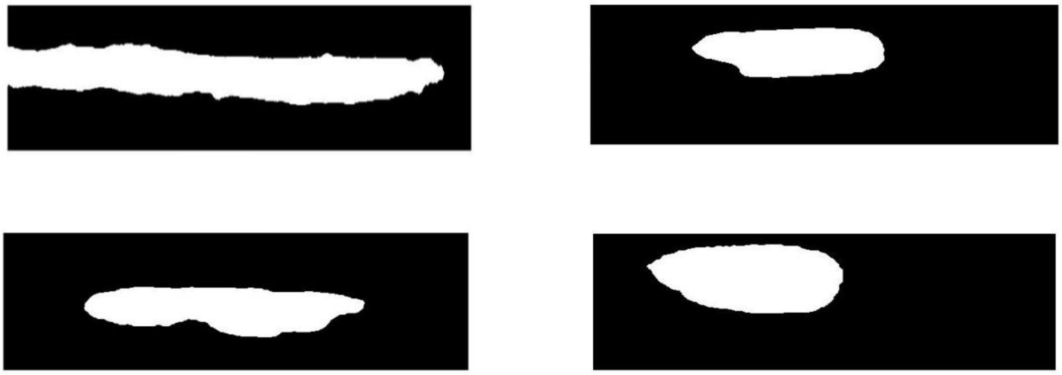

Fig 9: Binarized results of bacterial leaf blight, brown spot, narrow brown spot and rice blast respectively

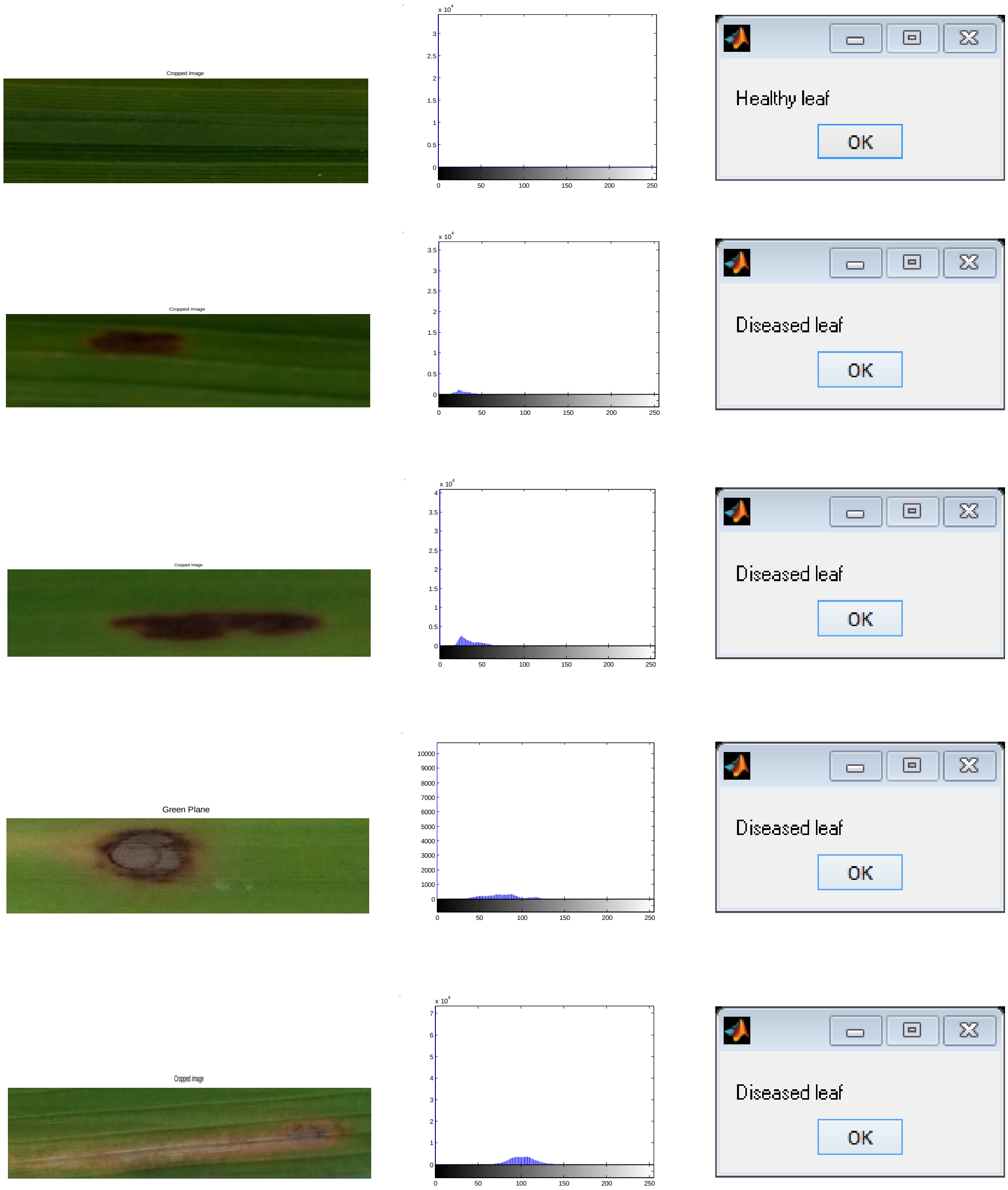

Figure 10:Original images and the Histogram plots of normal, bacterial leaf blight, brown spot, narrow brown spot and rice blast are given in (a)-(e), (f)-(j) respectively. And the results are given in (k)-(o)

The resultant images are used for the shape and color feature extraction. Classification of bacterial leaf blight, brown spot, narrow brown spot and rice blast diseases are carried out using SVM classifier. Here, 60 samples are used for classification purpose. For training 10 features are considered for each disease and for testing 5 samples for each disease. 70% accuracy has been achieved by using SVM classifier. TABLE 1 shows the resultant confusion matrix for classification of four diseases.

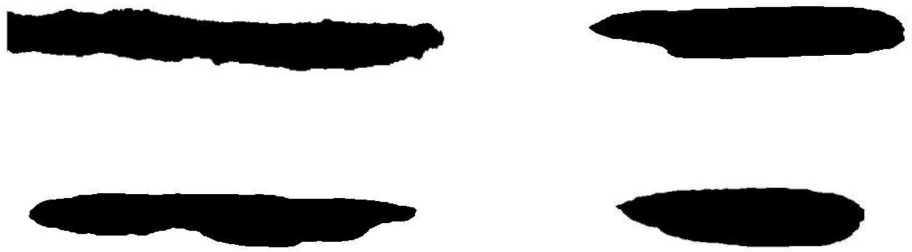

Fig 11: Segmented results of bacterial leaf blight, brown spot, narrow brown spot and rice blast respectively.

Table 1: Confusion Matrix for Classification

	Bacterial leaf blight	Brown spot	Narrow brown spot	Rice Balst
Bacterial leaf blight	5	0	0	0
Brown spot	2	2	0	1
Narrow brown spot	1	0	4	0
Rice Balst	2	0	0	3

VII CONCLUSION

Identification of the symptoms of plant diseases by means of image processing techniques is of prime concern in the area of research. An introduction to the research in agriculture field and different types of diseases in rice leaf is given. The literature survey done in preprocessing techniques and segmentation of leaf disease detection and classification has been discussed. After applying the suitable preprocessing technique classification of normal and diseased leaf using histogram plot is presented. Shape features are extracted using PCA method and the color features are extracted by using color based grid moments. These features are combined and fed to the SVM classifier. 70% accuracy is achieved for four different diseases.

REFFERENCES

[1] Srivastava K.P, *a text book of applied entomology* New Delhi: Kalyani Publishers.

[2] Rice Doctor, International Rice Research Institute, Philipines, www.irri.org, 2003.

[3] Phadikar.S, Sil. J, *Rice disease identification using pattern recognition techniques*, IEEE 11th International Conference on Computer and Information Technology2008. ICCIT.

[4] J.N. Kapur, P.K. Sahoo, A.K.C. Wong, *A new method for gray-level picture thresholding using the entropy of the histogram*, Computer Vision Graphics Image Process. *Vol. 29, pp.273-285, 1985*

[5] P. Sanyal, U. Bhattacharya and S. K. Bandyopadhyay ,*Color Texture Analysis of Rice Leaves Diagnosing Deficiency in the Balance of Mineral Levels Towards Improvement of Crop Productivity*, IEEE 10th International Conference on Information Technology, pp. *85-90, 2007.*

[6] Kurniawati, N.N. Abdullah, S.N.H.S. ; Abdullah, S. *Investigation on Image Processing Techniques for Diagnosing Paddy Diseases*, IEEE International Conference of Soft Computing and Pattern Recognition, *SOCPAR '09. 4-7 Dec. 2009*.

[7] K. J. Yoon and I.S. Kweon, *Color Image Segmentation Considering the Human Sensitivity for Color Pattern Variations*, Photonics Boston, *SPIE, 2001, pp. 269-278*.

[8] H. Tian, S.K. Lam and T. Srikanthan, 2003. *Implementing Otsus Thresholding Process using Area-Time Efficient Logarithmic Approximation Unit*, proc. *Of ISCAS03, IV-21-IV-24.*

[9] A.K. Jain, 1989. *fundamental of digital image processing*. NJ: Prentice Hall.

[10] Libo Liu and Guomin Zhou, *Extraction of the Rice Leaf Disease Image Based on BP Neural Network*. IEEE International Conference on Computational Intelligence and Software Engineering. *CiSE 2009*.

[11] Asfarian, A. Herdiyeni, Y.Rauf, A.Mutaqin, K.H, *Paddy diseases identification with texture analysis using fractal descriptors based on fourier spectrum*. IEEE Conference on Computer, Control, Informatics and Its Applications *(IC3INA), 2013* International .