2016 International Conference on Computing, Analytics and Security Trends (CAST)
College of Engineering Pune, India. Dec 19-21, 2016

Monitoring and Controlling Rice Diseases Using Image Processing Techniques

Amrita A. Joshi
E &Tc Department
JSPM's RSCOE
Savitribai Phule Pune University
Pune, India
joshiamruta07@gmail.com

B.D. Jadhav
E &Tc Department
JSPM's RSCOE
Savitribai Phule Pune University
Pune, India
bdjadhav1979@gmail.com

Abstract—**Development of an automated system for identifying and classifying different diseases of the contaminated plants is an emerging research area in precision agriculture. Identification of the diseases is the key to prevent qualitative and quantitative loss of agricultural yields. Rice (Oryza Sative) is one of the essential crops in India and losses due to the diseases badly impact the economy. Manual detection of the diseases is very difficult and not accurate. This creates a need for Image processing techniques which will help in accurate and timely detection of the diseases and overcome the limitations of the human vision. A new technique to diagnose and classify rice diseases has been proposed in this paper. Four diseases namely rice bacterial blight, rice blast, rice brown spot and rice sheath rot have been identified and classified. Different features like shape, the color of a diseased portion of the leaf have been extracted by developing an algorithm. All the extracted features have been combined as per the diseases and diseases have been classified using Minimum Distance Classifier (MDC) and k-Nearest Neighbor classifier (k-NN). The performance of the proposed technique has been evaluated with the help of 115 rice leaf images of four diseases and 70 percent image data has been used for training the classifier and 30 percent has been used for testing. Classification accuracy has been calculated for each disease using both classifiers. The overall accuracy achieved by using k-NN and MDC is 87.02 percent and 89.23 percent respectively.**

Keywords—***k-Nearest Neighbor classifiers, Minimum Distance Classifier, Regional properties, Support Vector Machine, YCbCr color space.***

I. INTRODUCTION

Rice (Oryza Sativa) is one of the important crops in India. The majority of the land area is under the cultivation of the rice crop and is one of the main sources of economic development. Every year farmers face a loss of yield and financial losses due to the pests and the diseases on the rice plants. Bacteria, fungi, and viruses are the main reasons of the rice diseases. There are many rice diseases, out of which four diseases have been studied in this paper namely Rice Bacterial Blight (RBB), Rice Blast (RB), Rice Brown Spot (RBS) and Rice Sheath Rot (RSR). These diseases have some similar symptoms which confuse the human vision while detecting them. Early symptoms of some diseases cannot be detected by the farmers. Farmers require continuous monitoring by experts in order to diagnose diseases at an early stage which are very expensive and time-consuming. This creates a need for image processing techniques to detect diseases automatically. Image processing techniques help not only to diagnose the diseases accurately but also to classify diseases. Early and accurate identification of the diseases is also possible. This will increase the productivity and quality of the rice yield. With image processing techniques, human efforts are reduced in great extent. Image processing techniques are widely used for different agricultural applications ranging from identification of the leaflet of the plant to categorization of different diseases. In general, image processing techniques have following steps to identify the diseases: 1)Image acquisition 2) Image preprocessing 3) Segmentation 4)Feature Extraction 5)Classification. Different methods have been used for segmentation, feature extraction and for classification various classifiers are available.

Anthonys G. and N. Wickramarachchi proposed three rice diseases detection technique, rice blast, rice sheath blight and brown spot. After image acquisition by a digital camera, RGB image was digitized and segmentation of the image was done using mathematical morphology. Features like shape, texture, and color were extracted and used as an input for the classification. Extracted three features were used in a method of membership function to distinguish between the above-mentioned diseases. They achieved over 70% classification accuracy for 50 sample images. But preprocessing is required to reduce the noise effect [1].

Segmentation using threshold value was applied by Kurniawati, Nunik Noviana, et al. To remove the noise they used region filling technique. The shape and color features were extracted using texture analysis and the classification was done using production rule [2]. Hue Intensity Saturation (HIS) model for segmentation, Boundary detection algorithm using the 8-connectivy method and classification using self-organizing map neural network was proposed by Phadikar et al [3]. On the basis of morphological changes, two diseases brown spot and leaf blast were classified. Otsu’s threshold based segmentation, Bayer's theorem, and SVM were used to classify the diseases[4].

Cheng and F.Y. Shih presented an automatically seeded region growing algorithm for color image segmentation. YCbCr color space is selected to transform color image. The initial seeds were automatically selected, the color image was segmented into regions where each region corresponds to a seed and region-merging was used to merge similar regions [5]. Sarkar, Rajat Kanti, and Ankita Pramanik, modified an automatically seeded region growing (SRG) algorithm proposed by Cheng and F.Y. Shih. Euclidean distance was calculated to get the color difference between adjacent regions. Region merging was done using two-dimensional look-up table and it was then used for segmentation. To check the performance of segmentation algorithm different color spaces were used. Algorithm showed best segmentation results in YCbCr space [6]. Color and different shape features were extracted by Phadikar S. and Jaya Sil. To improve the simplicity of the algorithm, rough set theory (RST) was implemented and features were selected [7].

Back propagation neural network (BPNN) was used by L. Liu and G. Zhou to check whether the rice leaves were healthy or infected [8]. Kholis Majid, Yeni Herdiyeni, Aunu Rauf implemented fuzzy entropy to extract features and probabilistic neural network (PNN) to classify the four different diseases and 5-fold cross-validation has been used for testing [9],[13].

This paper is arranged as:
Section II. Symptoms of the four diseases, III. Proposed methodology, IV Experimental results and finally in section V concluded the paper.

II. Symptoms of Diseases

Symptoms of diseases are classified as given below:

A. *Rice Bacterial Blight(RBB)*

Bacterial blight is created by *Xanthomonas Oryza pm. Oryza*. Seedling becomes limp, water-soaked to yellowish stripes on a leaf or beginning at leaf tips then later raise in length and width with a wavy margin. Lesions turn yellow to white as the disease progresses [14], as shown in Fig.1(a).

B. *Rice Blast(RB)*

It is because of fungus *Magnaporthe Oryza.* White to gray-green lesions or spots, with dark green borders in an initial stage. More established injuries on the leaves are elliptical or spindle-shaped and whitish to gray centers with red to a brownish or necrotic border, shown in Fig.1(b). Spots are usually lengthened and pointed at each end [14].

C. *Rice Brown Spot(RBS)*

It is created by the fungus *Sphaerulina oryzina* (syn. *Cercospora Jan Sena*, *Cercospora oryza*) .The typical wound on leaves and upper leaves are light to dark brown, straight, and progress parallel to the vein. They are usually 2−10 millimeter long and 1−1.5 millimeter wide as shown in Fig.1(c). Injuries on the leaves of highly susceptible varieties may enlarge and connect together that forming brown linear necrotic regions.
Brown lesions are also found on pedicels. The disease also causes staining on the leaf sheath, alluded to as “net blotch” because of the netlike pattern of brown and light brown to yellow areas. Thin brown spot can be mistaken for white leaf streak. Linear lesions make the disease distinct from other leaf diseases [14].

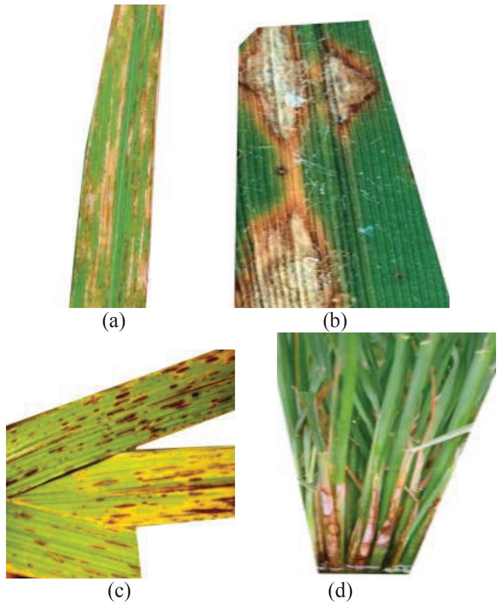


Fig.1 Images showing symptoms of rice diseases (a) RBB, (b) RB, (c) RBS and (d) RSR

D. *Rice Sheath Rot(RSR)*

It is created by two fungal species, *Sarocladium Oryza and Sacroladium attenuatum*. The typical sheath rot injury starts at the uppermost leaf sheath encasing the little panicles. It seems oblong or as an asymmetrical spot with dark reddish, brown margins, and gray centers or brownish gray all through. Usually, several spots are observed and these spots enlarge and merge or raise together and can cover up most of the leaf sheath. Panicles remain within the sheath or may partially emerge. Affected leaflets may have plentiful whitish powdery fungal growth (mycelium) visible on the outer surface. Panicles that have not appeared rot and the florets turn red-brown to dark brown [14]. The image is shown in Fig.1(d).

III. Proposed Methodology

The methodology for identifying and classifying rice diseases involves pre-processing, segmentation, feature extraction, classification. Fig.2 shows the training phase of the proposed methodology. In training phase, few images have been trained by extracting their feature values and storing them in a file. This trained dataset is further used as reference data for the classifier to classify the input or test images. Fig.3 shows the testing phase of the proposed methodology. In the testing phase, unknown image or test image is processed and its features are given as input to the classifier. The classifier classifies them using the features stored in dataset during the training phase.

A. *Data collection and Preprocessing*

In this work, JPEG images of four infected rice leaves were collected. Some of these images are taken from Agricultural Research Station, Lonavala, Maharashtra. Unfortunately, there are only a few samples available for these diseases. To compensate this condition, we took some other images from the Internet that have different size and quality than our images.

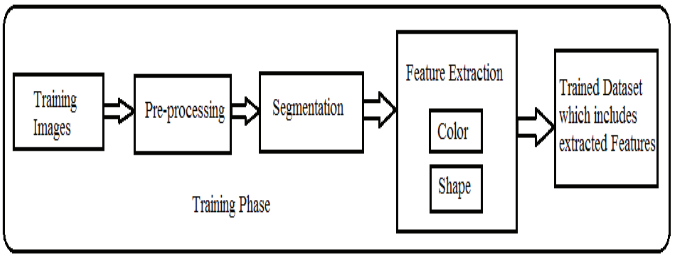


Fig.2 Training phase of proposed methodology

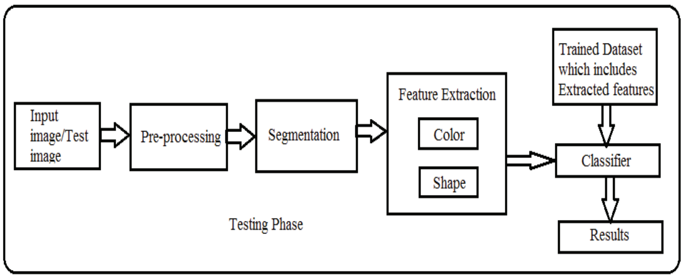


Fig.3 Testing phase of proposed methodology

Total 115 images of sick paddy leaves are collected. The input image is first resized to 200 X 200 pixels. In order to remove the effect of outdoor illumination, normalize the image. To normalize the image, extract R,G,B components and take the mean of three components and find the maximum value out this mean. By dividing each extracted mean by this maximum value gives the scaling factor for each R,G,B plane. Multiply the R,G,B component values with this scaling factor, which results in the normalized image.

B. Segmentation

Convert the normalized RGB image into YCbCr space. YCbCr color space has been used mainly for segmentation of rice diseases because of below reasons 1) In YCbCr plane infected part can be easily detected 2) The color difference of human discrimination can be directly uttered by Euclidean distance in the YCbCr color space 3) The intensity and chromatic components can be used separately and 4) Plant infected spots form small clusters in Cr space [5]. Values of Y, Cb, Cr can be calculated using (1) - (3)

$$Y = 0.2999R + 0.587G + 0.114B + 16 \quad (1)$$

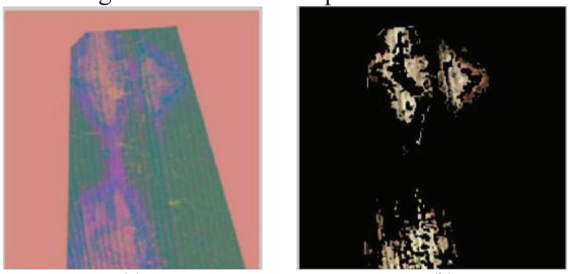
$$C_b = -0.169R - 0.331G + 0.500B + 128 \quad (2)$$

$$C_b = -0.169R - 0.331G + 0.500B + 128 \quad (3)$$

By observing different images in YCbCr plane , Cb and Cr range for a diseased part is determined. After calculation of the range of Cb and Cr, check the row and column number of the image which satisfy this range and convert that rows and columns pixels to white (assign the value 255 to a pixel). Convert the image into binary and multiply this with the original RGB image, this will result in the segmented image.

C. Feature Extraction

When rice leaves are infected by any disease, color of the leaves is changed and the colored spots are created on leaves.



(a) (b)

Fig.4 (a) RGB image converted into YCbCr space (b) Segmented image

These spots are of different shapes. Hence, color and shape of the infected part are selected as features to identify the diseases.

I) Color Feature Extraction

The color of the diseased part of the leaf is one of the important indication for detecting the different diseases. The color of the affected part is different for every disease and also varies for different varieties of the rice plant. The color of the diseased part for the same disease is also different. R, G, B components of the segmented image are extracted. Color moments like mean and standard deviation are considered as the color features. Mean and standard deviation in three R,G,B plane are calculated. Total 6 color features are extracted. Value of mean is calculated using (4) and value of standard deviation is calculated using (5)

$$E_i = \sum_{j=1}^{N} \frac{1}{N} p_{ij} \quad (4)$$

where,

N is number of pixels in an image p_{ij} is j^{th} pixel of image at i^{th} color channel

$$\sigma_i = \sqrt{\left(\frac{1}{N}\sum_{j=1}^{N}\left(p_{ij} - E_i\right)^2\right)} \quad (5)$$

where,

E_{ij} is the mean value or first color moment, for the i^{th} color channel of the image.

II) Zone Wise Shape Feature Extraction

When plants are infected by pathogens they create spots of different shapes. Shapes can vary as per the type of pathogen, type of rice species, and type of diseases. Irregular, elliptical, oval, rectangular, circular shapes are created by pathogens. In image processing techniques, regional properties have been used to describe and represent the shapes. The segmented image is converted into a binary image and the number of connected components is calculated. Area and centroid of each connected component are calculated. The area is the number of white pixels in a given image and centroid is the center of mass of the given region. To reduce the time required for the extracting shape of each component, the component having maximum area is cropped as shown in Fig.5(a) and (b) and used for further processing. The cropped region of the infected part has to be represented and described. Morphological operation 'skel' is performed on the cropped area of an infected part.

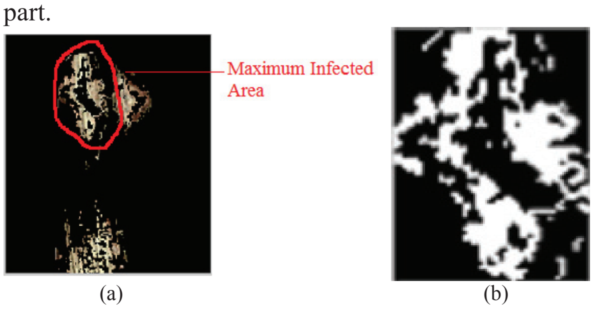


Fig.5 Cropped maximum infected area

Image after 'skel' operation is as shown in Fig.6 (a).This operation prepares the skeleton of the infected cropped part. Shape descriptors like eccentricity, extent, and orientation are considered as shape features and calculated for a given image, Fig.6 (a). Eccentricity is the ratio of the distance between the foci of the ellipse and its major axis length. Extent that specifies the ratio of pixels in the region to pixels in the total bounding box and orientation is the angle (in degrees ranging from -90 to 90 degrees) between the x-axis and the major axis of the ellipse [16].

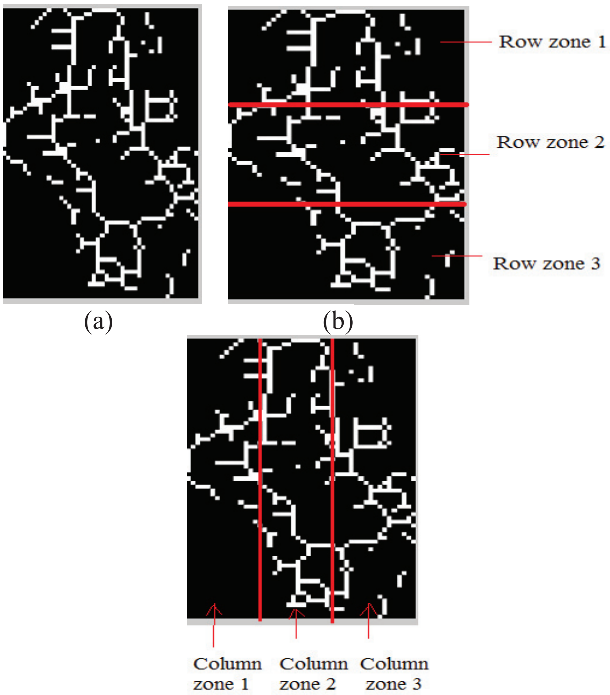


Fig.6. (a) Skeleton of the cropped area (b) Showing three row zone (c) Showing three column zone

Cropped image is divided into three equal size row zone, shown in Fig.6(b) and three equal size column zone Fig.6(c). Column zone height, column zone width , row zone height and row zone width of each zone is calculated using following equations

$$\text{Row zone height} = \text{No.of rows}/3 \tag{6}$$

$$\text{Row zone width} = \text{No. of columns} \tag{7}$$

$$\text{Column zone height} = \text{No. of rows} \tag{8}$$

$$\text{Column zone width} = \text{No. of columns}/3 \tag{9}$$

From height and width of both row and column zone, the size of each zone has been calculated. For each row zone and column zone, the starter points and intersections are calculated. All the pixels in the given image have been checked to decide whether it is a starter point or intersection point. Starter points are those having only one neighbor and intersections are those pixels having three or more neighbors. Depending on different conditions of the starter and intersection points a number of line segments in each zone and their direction is calculated.The logic to find the line segment direction is similar to the 8 direction chain codes. From these segment and segment direction, following nine features are extracted from each row zone and column zone.

The logic to find following features is based on feature extraction technique used for recognition of hand written character [10].

1) The no. of horizontal lines.
2) The average length of horizontal lines.
3) The no. of right diagonal lines.
4) The average length of right diagonal lines.
5) The no. of vertical lines.
6) The average length of vertical lines.
7) The no. of left diagonal lines.
8) The average length of left diagonal lines.
9) The no. of intersection points.

All the extracted features are combined and used as an input to classifiers.

D. *Classification*

Two classifiers namely MDC and k-NN has been applied for the classification of diseases. Distance classifier is preferable because of its simplicity of usage and less classification time [11,12]. MDC with Manhattan distance and k-NN with the 'cityblock' distance metric are implemented. To test the performance of the proposed technique, 70 % data is used for training purpose and 30% data is used for testing purpose. Manhattan distance is also called cityblock distance and is a distance (d), between two points A(x_1, y_1) and B(x_2, y_2) , given by (10).

$$d = |x_1 - x_2| + |y_1 - y_2| \tag{10}$$

E. *Remedies For Diseases*

Table I gives the remedies for particular rice disease which are commonly used by farmers. Out of these , few remedies has been suggested by the rice pathologist

IV. Experimental Results

The proposed method is implemented using MATLAB R2012a version, with image processing toolbox and tested on the personal computer with Windows 10 operating system. The database used contains 115 images of four diseases, out of which 70% data has been used for training the classifier and 30% data has been used for testing. Pre-processing, segmentation, feature extraction and classification steps have been carried out for each image. Confusion matrix for MDC and k-NN is given in Table II and III respectively, and classification accuracy is calculated from it. Classification accuracies for four diseases are calculated using both the classifiers. Classification accuracy of both classifiers is same for two diseases namely rice blast and rice brown spot whereas for other two diseases namely blight and sheath rot; classification accuracy of MDC is more than k-NN classifier. Classification accuracy is the ratio of a number of correctly identified images and a total number of images, given in (11).

$$\text{Classification Accuracy}=\frac{\text{No. of correctly classified images}}{\text{Total no. of images}} \qquad (11)$$

TABLE I. REMEDIES FOR DISEASES

Disease Type	Remedies
RBB	Spray crop with Streptomycin sulphate (200ppm) + Copper oxychloride (0.25%) + sticker (0.10%). First spray immediately after disease appearance followed by 1 to 2 sprays at 10 to 15 days interval
RB	Carbendazim 50WP @ 500 grams/hectare Tricyclozole 75 WP @ 500 grams/ hectare Metominostrobin 20 SC @ 500 milliliters / hectare Azoxystrobin 25 SC @ 500 milliliters / hectare Propiconazole 500 milliliters / hectare [15]
RBS	Seed treatment with iprodione (0.25%) or carbendazim (0.10%)
	Use fungicides *viz.*, iprodione (0.25%), propiconazole (0.10%), azoxystrobin(0.05%), trifloxystrobin (0.10%), chlorothalonil (0.25%), Metominostrobin (0.10%), etc. for foliar sprays [15]
RSR	Seed treatment with fungicide like Carbendazim or benomyl @ 3.0 g/kg seed.
	Apply a foliar fungicide like Carbendazim or propiconazole or Metominostrobin or Hexaconazole 75% WG @ 500grams/hectare [15]

TABLE II. CONFUSION MATRIX FOR MDC

Disease types determined from classifier	**Disease types determined from reference source**				
	RBB	***RB***	***RBS***	***RSR***	***Total***
RBB	23	1	1	1	26
RB	3	29	1	0	32
RBS	0	2	29	1	31
RSR	2	0	0	18	20

TABLE III. CONFUSION MATRIX FOR K-NN

Disease types determined from classifier	**Disease types determined from reference source**				
	RBB	***RB***	***RBS***	***RSR***	***Total***
RBB	22	1	1	2	26
RB	3	29	1	0	32
RBS	0	1	29	2	31
RSR	2	1	0	17	20

Table IV and Fig.7 shows the classification accuracy of the classifiers of four diseases. Fig.7 shows that the proposed technique gives better classification results for disease brown spot and sheath rot with MDC classifier.

TABLE IV. CLASSIFICATION ACCURACY OF CLASSIFIERS

Classifiers / Disease Type	**Minimum Distance Classifier**	**k-Nearest Neighbor**
RBB	88.46%	84.61%
RB	87.87%	87.87%
RBS	90.62%	90.62%
RSR	90%	85%

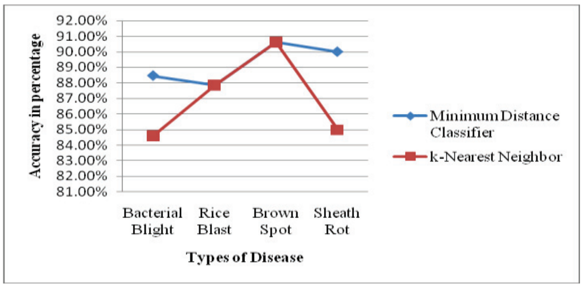


Fig.7 Classification accuracy of proposed technique

The proposed technique is compared with existing techniques and observed that accuracy of the proposed techniques has been improved up to 89.23%. Table IV gives the comparison of proposed technique and the existing techniques

TABLE V. COMPARISON OF CLASSIFICATION ACCURACY

Techniques	**Classification accuracy**
Anthonys G., & N. Wickramarachchi[1]	70 %
Kurniawati, Nunik Noviana, et al[2]	86 %
Phadikar S.& J. Sil[3]	82%
Phadikar S. & J. Sil[3] [4]	68.10%
Proposed technique	89.23 %

V. CONCLUSION AND FUTURE SCOPE

A system for monitoring rice bacterial blight, rice blast, rice sheath rot and brown spot diseases has been developed in this study. Image processing techniques like segmentation, feature extraction, and two classifiers were used to establish the classification algorithm.

In the feature extraction, color and zone wise shape features have been extracted and used as an input to the classifier. For each disease, a separate database has been used for training and testing. Proposed techniques have been tested for four mentioned rice diseases using two classifiers, k-NN and MDC and the accuracy achieved with two classifiers is 87.02 % , 89.23% respectively. The proposed technique is compared with the few existing techniques which are related to the rice diseases detection and found that the proposed technique is superior in terms of time complexity, accuracy, number of diseases covered.

In this work, color and shape features extraction have been carried out. In addition to this, texture feature can be included and can be checked for the impact of this extra feature on the performance of an algorithm**.** There are other rice diseases except four covered in this work. Future work can be to cover other rice diseases. The same techniques can be applied to other crops with little modifications.

ACKNOWLEDGMENT

We thank Dr. A. P. Gaikwad, Rice Pathologist, Agricultural Research Station, Lonavala, for providing the rice diseases database

REFERENCES

[1] Anthonys G., and N. Wickramarachchi, "An image recognition system for crop disease identification of paddy fields in Sri Lanka," In *Industrial and Information Systems (ICIIS), International Conference on*, pp. 403-407, 2009.

[2] Kurniawati, Nunik Noviana, and Saad Abdullah,"Texture analysis for diagnosing paddy disease," In *Electrical Engineering and Informatics, International Conference on*, vol. 1, pp. 23-27, IEEE, 2009.

[3] S.Phadikar and J. Sil, "Rice disease identification using pattern recognition techniques," *Computer and Information Technology,11th International Conference on*, Khulna, pp. 420-423, 2008.

[4] S.Phadikar, J. Sil, and A. K. Das, "Classification of Rice Leaf Diseases Based on Morphological Changes," International Journal of Information and Electronics Engineering vol. 2, no. 3, pp. 460, 2012.

[5] F.Y. Shih and S. Cheng, "Automatic seeded region growing for color image segmentation," *Image and Vision Computing*, Vol. 23, pp. 877-886, 2005.

[6] Sarkar Rajat Kanti and Ankita Pramanik, "Segmentation of plant disease spots using automatic SRG algorithm: a look-up table approach," In Computer Engineering and Applications (ICACEA), 2015 International Conference on Advances in, pp. 1-5, IEEE, 2015.

[7] S. Phadikar , J. Sil and A. K. Das, "Rice diseases classification using feature selection and rule generation techniques,"*Computers and electronics in agriculture,* vol. 90, pp. 76-85, 2013

[8] L. Liu and G. Zhou, "Extraction of the Rice Leaf Disease Image Based on BP Neural Network," *Computational Intelligence and Software Engineering, 2009, CiSE 2009. International Conference on*, Wuhan, pp. 1-3, 2009

[9] K. Majid, Y. Herdiyeni and A. Rauf, "I-PEDIA: Mobile application for paddy disease identification using fuzzy entropy and probabilistic neural network," *Advanced Computer Science and Information Systems (ICACSIS), 2013 International Conference on*, Bali, pp. 403-406, 2013.

[10] M. Blumenstein, B. Verma, and H. Basli, "A novel feature extraction technique for the recognition of segmented handwritten characters," *Document Analysis and Recognition, 2003. Proceedings. Seventh International Conference on*, vol.1, pp. 137-141, 2003.

[11] Golubitsky Oleg, and Stephen M. Watt, "Distance-based classification of handwritten symbols," *International Journal on Document Analysis and Recognition* (IJDAR) 13, no.2, pp.133-146, 2010 .

[12] Junhua Li and Li Peng, "Human expression recognition based on feature block 2DPCA and Manhattan distance classifier," *Intelligent Control and Automation, WCICA 2, 7th World Congress on*, Chongqing, pp. 5941-5945, *2008.*

[13] Amrita A.Joshi, and B.D.Jadhav, "Survey On Role Of Image Processing Techniques in Plants Pathology ," National Conference On Innovation in Engineering Science and Technology, pp. 135-139, 2015.

[14] http://www.knowledgebank.irri.org

[15] http://agritech.tnau.ac.in/crop_protection/crop_prot_crop%20diseases_cereals_paddy.html

[16] http://in.mathworks.com/help/images/ref/regionprops.html