A Survey on Detection and Classification of Rice Plant Diseases

Jitesh P. Shah
Department of Information Technology
Dharmsinh Desai University, Nadiad, Gujarat, India
jitesh2k12@gmail.com

Harshadkumar B. Prajapati
Department of Information Technology
Dharmsinh Desai University, Nadiad, Gujarat, India
harshad.b.prajapati@gmail.com

Vipul K. Dabhi
Department of Information Technology
Dharmsinh Desai University, Nadiad, Gujarat, India
vipul.k.dabhi@gmail.com

***Abstract*—Identifying disease from the images of the plant is one of the interesting research areas in computer and agriculture field. This paper presents a survey of different image processing and machine-learning techniques used in the identification of rice plant diseases based on images of disease infected rice plants. This paper presents not only survey of various techniques but also concisely discusses important concepts of image processing and machine learning applied to plant disease detection and classification. We carry out detailed study of 19 papers, covering the work on rice plant diseases and other different plants and fruits, and present a survey of these papers based on important criteria. These criteria include size of image dataset, no. of classes(diseases), preprocessing, segmentation techniques, types of classifiers, accuracy of classifiers etc. We utilize our survey and study to propose and design our work on detection and classification of rice plant diseases.**

Keywords—image processing; machine learning; classification; clustering; disease classification; disease detection.

I. Introduction

Agriculture is the main source of the income in India. Almost 70 percent of the population in India depends on farming[1]. Over 58 percent of the rural people depend on farming as it is an essential mean of livelihood [2]. Rice is one of the important foods in India. Rice disease destroys 10 to 15 % of production in Asia [3]. Fungus, bacteria are responsible for disease in the plant. Different diseases that occur on rice plants are Leaf blast, Brown spot, Sheath blight and Leaf scald[4].

Sometimes farmers are unable to pay attention to the diseases or face difficulty in identifying the diseases, which lead to loss of the crop. Every disease has a different remedy to work out. For example, fungi based disease can be prevented by disrupting the life cycle of the pathogen[4]. The current approach of disease detection is manual, which means farmers mainly depend on the guide books or use their experiences to identify the diseases[5]. Each plant disease has different stages of growth. Whenever the disease occurs on a plant, farmers have to keep eyes on the infection. This approach of disease detection is time-consuming and requires some precaution during the selection of pesticides.

Capturing images of "seem to appear" infected leaves and getting them processed by an automated system could become an attractive solution for farmers. For example, an automated system can be deployed in the farm, in which camera sensors put at different location in the farm captures images periodically and the system processes captured images for any infection. Using such system, farmers can be informed about diseases instantly. This paper focuses on how image processing and machine learning is utilized in detection of diseases in rice plants. The general process of plant disease identification is shown in Fig.1. Major steps of the processing are as follows. The images are captured from the field, the images are preprocessed, the infected parts are extracted from the leaf, then features are extracted from the segmented images, and finally classification of the disease is performed using machine learning techniques.

Success of such system depends on how accurately the system carries out both needed image processing and machine learning operations. The performance of a plant disease detection system can be evaluated by measuring the accuracy of the machine learning algorithms (classification). Existing systems cannot provide better accuracy due to limited image features, bad selection of classifiers, and bad selection of features. Moreover, problems like shadow effect in images may lead to misclassification of disease [6]. Existing systems may not be feasible due to different diseases having similar spots and a same disease may have different spots because of different rice varieties and local conditions[5].

Due to possibility of various alternatives at different states of plant disease identification, which is shown in Fig.1, researchers have attempted various alternatives in both image processing operations and machine learning models. To get comparative understanding of various works, we carry out a survey on different techniques and approaches used for the detection of the Rice plant diseases. To get understanding of the domain, we carry out a study on different types of rice plant diseases. We survey both image processing techniques and machine learning techniques applied in Rice disease classification and present our findings derived from our survey. Furthermore, utility of the presented survey and findings is shown in our proposed work in the same direction.

This paper is divided into eight sections. Section II presents study on different types of rice plant diseases. Section III presents the process of plant disease identification and discusses various operations. Section IV presents research that has already been done on plant disease detection. Section V presents a survey of different image processing operations

978-1-5090-1936-6/16/$31.00 ©2016 IEEE

applied for rice disease identification and identifies important findings. Section VI presents a survey of machine learning operations applied for rice disease identification and identifies important findings. Section VII uses the survey and findings to prepare our proposal about detection and identification of diseases in rice plant. Finally, Section VIII summarizes work in the form of conclusion.

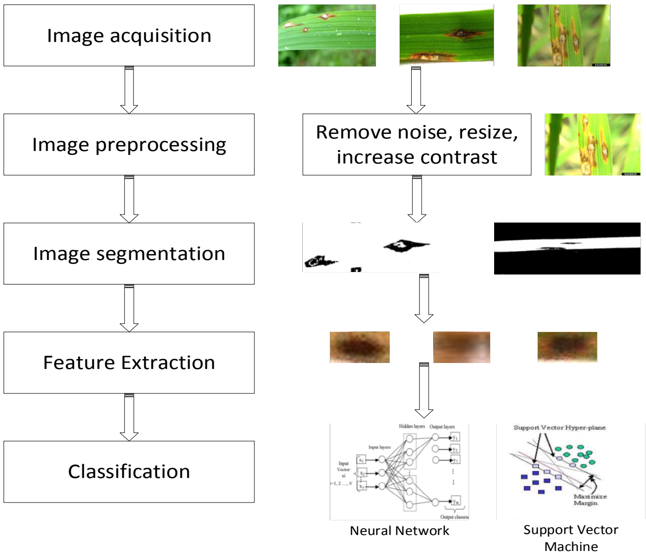

Fig. 1. General Approach of Plant Disease Identification from the Images

II. DIFFERENT TYPES OF RICE PLANT DISEASES

This section briefly explains different types of diseases that occur on rice plants. The reason behind putting this section is that the readers can understand about what types of image processing operations would be needed and what type of features need to be considered to prepare such disease detection system. The images of six frequently occurring diseases are shown in Fig.2. We describe each disease briefly. Further details on all types of rice plant diseases can be found in [7].

1) Leaf blast: The symptoms of this disease are dark spot to oval spot with narrow reddish-brown margins and gray or white center[7].

2) Brown spot: This disease occurs on leaves of the rice plant. The symptoms of the disease are round to oval shape with dark brown lesions[7].

3) Sheath blight: This disease occurs on both leaves and stems. The symptoms are oval, white or straw colored regions in the center with reddish brown spots[7].

4) Leaf scald: The symptoms are narrow reddish-brown wide bands. Sometimes lesions are at leaf edges with yellow or golden borders[7].

5) Bacterial leaf blight: Symptoms contain elongated lesions on leaf tip , lesions are several inches long and it turn into yellow from white due to effect of bacteria[7].

III. GENERAL PROCESS OF PLANT DISEASE IDENTIFICATION

This section explains the General process of plant disease identification. The process is divided into two parts: (1) Image processing and (2) Machine learning.

A. *Image Processing Tasks*

1) *Image Acquisition*

An image database specifically for rice disease pictures is available at IRRI (International Rice Research Institute)[8]. Therefore, we need to prepare image database by our own, which requires image acquisition from live farm. In this process, images are captured from the farm using a digital camera to get them directly in digital form with numerical values[9] so that digital image processing operations can be applied. Some of widely used digital cameras for the purpose are Nikon D80[5], Nikon Coolpix P4[6], Canon 450D[10], Canon PowershotG2[11] etc.

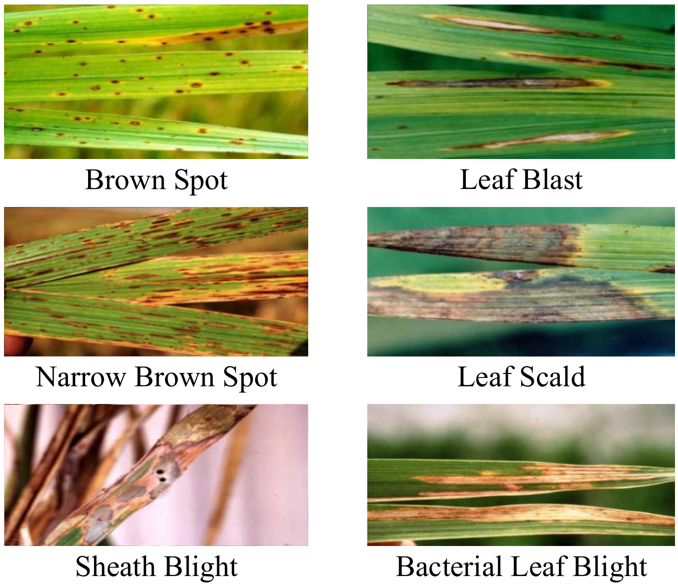

Fig. 2. Different Types of Rice Diseases. Source of Images[7]

2) *Image Preprocessing*

For getting better results in further steps, image preprocessing is required because dust, dewdrops, insect's excrements may be present on the plant; these things are considered as image noise [5]. Furthermore, captured images may have distortion of some water drops and shadow effect, which could create problems in the segmentation and feature extraction [5] stages. Effect of such distortion can be weakened or removed using different noise removal filters. There may be low contrast in captured images; for such images contrast enhancement algorithms can be used. Sometimes background removal techniques may also be needed in case of region of interest needs to be extracted. In case of noise such as salt and pepper, a median filter can be used. Weiner filter can be used to remove a blurring effect. In case of the images captured using high definition cameras, the size of the pictures might be very large, for that reduction of image size is required. Also, image reduction helps in reducing the computing memory power [9].

3) *Image Segmentation*

Image segmentation can play a vital and important role in plant disease detection. Image segmentation means to divide the image into particular regions or objects. The primary aim of segmentation is to analyze the image data so one can extract the useful features from the data. There are two ways to carry out the image segmentation: (1) based on discontinuities and (2) based on similarities. In the first way, an image is partitioned based on sudden changes in intensity values, e.g., done via edge detection. While in the second way, images are partition based on the specific predefined criteria, e.g., thresholding done using Otsu's method.

4) *Feature Extraction*

The feature extraction aspect of image analysis focuses on identifying inherent characteristics or features of objects present within an image. These features can be used to describe the object. Generally, features under following three categories are extracted: color, shape, and texture. The color is an important feature because it can differentiate one disease from another. Furthermore, each disease may have different shape; thus system can differentiate diseases using shape features. Some shape features are area, axis, and angle. Texture means how color patterns are scattered in the image [22].

B. *Machine Learning Tasks*

1) *Classification*

Classification maps the data into specific groups or classes. Classification is usually called as supervised learning approach [12]. Classification is a two-step process: First the classifier model is generated which describes predefined set of classes. This step is called as learning phase (Training step), where classification algorithm develops the classifier by "learning from" the data with their specific class labels. In the second step, the model, which is generated in first step, is used for classification [13]. In other words, test data is used to estimate the accuracy of the trained model by evaluating how good it performs on the test data. In the plant disease classification, the diseases are classified according to the features extracted from the images. Different classification models are support vector machine, neural network, nearest neighbour, and rule-based classifier.

2) *Clustering*

Clustering is a process of grouping data into different groups based on the similarity of the data. It means the data points with the similar objects are grouped into one group and dissimilar objects are grouped into another group [14]. Clustering is also called as data segmentation because it partitions large data into groups based on their similarities. Clustering is an unsupervised learning approach. Unlike classification, clustering does not depend on the predetermined classes, due to this clustering is called as learning by observation not learning by examples [13]. Clustering is also used for color image segmentation, because in an image different objects may have different color intensities. Therefore, clustering can group similar intensity pixels in one cluster and other different intensity pixels into other clusters.

IV. RELATED WORK

This section describes different works that have been already done by researchers in different domains such as fruit grading system, weed detection, classification of plants, fruits disease classification, etc.

Jayme Garcia et al.[15] have surveyed on the detection and classification of plant diseases. They also discussed on severity of the diseases. Their paper briefly described different image processing and machine learning techniques such as thresholding, neural networks, dual segmented regression analysis, quantification color analysis, fuzzy logic, and region growing. Rashmi Pandey et al. [16] reviewed image processing and machine learning techniques applied to the fruit grading system. Grading the fruits means according to size, shape, color, texture, stem, and calyx of fruits, the fruits are sorted. The color is the most important feature in an image. The paper briefly described different color feature extraction techniques. The paper also discussed a comparative study of the machine learning methods such as rule-based system, support vector machine, nearest neighbour, and artificial neural network.

Anup Vibhute et al. [17] had done a survey on different applications of image processing in agriculture. They made a survey on two different applications: weed detection and fruit/food grading system. Weeds are harmful to the crops, so from the images of crops, weed can be detected by different image processing techniques such as edge and color detection, classification based on wavelets and fuzzy, etc. Grading of fruits with different image processing concepts such as segmentation and feature extraction is discussed in that paper.

The work in [18] presented a study on different image processing and machine learning techniques used for detection of plant diseases using leaf images. First Color transformation structure is created, and then device independent color space transformation is applied on that color transformation structure. Pre-processing methods such as clipping, image smoothing, image enhancement, and histogram equalization are presented. Segmentation techniques such as boundary detection and spot detection, K-means clustering, and Otsu's thresholding are also briefly described. Key features such as Color, texture, morphology, and edges are also discussed. Classification based on ANN with the working principle is discussed.

The work in [19] presented plant disease detection with samples of two plants namely grape and wheat. Total 185 images including 85 grape leaves and 100 wheat leaves were considered. Without changing the image resolution, the images are compressed using nearest neighbour interpolation method. The median filter is applied to de-noise the images. K-means clustering is used for segmentation. For image segmentation, RGB image is converted into CIE XYZ Color space. After that XYZ color space is converted into l*a*b color space. Color, shape, and texture features are extracted. Principal component analysis (PCA) is used to reduce the dimensions of the feature data extracted from images. Due to this reduction in dimensions, the number of neurons can be reduced so that speed of Backpropagation neural network can increase.

The work in [20] proposed a work on detection and classification of plant diseases. Five diseases are considered:

Early scorch, Cottony mold, Ashen mold, Late scorch, and Tiny whiteness. The RGB images are converted into device independent CIE L*a*b* color transformation structure. K-means clustering is applied to extract diseased portion from leaf. Green colored pixels are detected and then it is masked in original image. The infected regions (clusters) are converted into HSI color space for extracting features. Spatial grey level dependency matrices are generated from hue and saturation plane of HSI model. Finally, neural network is used for classification.

V. Survey and Analysis of Image Processing Operations Applied for Rice Disease Identification

This section describes image processing techniques used in various works on rice disease detection. This section presents a survey of 13 papers of rice disease detection including criteria such as image dataset, no. of diseases, Preprocessing, Segmentation, Edge Detection, Feature Extraction, Image Background, and Color Space.

Table I presents the analysis of different papers of rice diseases identification. We use following notations in the table: Rice Blast(RB), Sheath Blight (SB), Brown Spot(BS), Bacterial Leaf Blight (BLB), Sheath Rot(SR), Narrow Brown spot (NBS), Red (R), Green (G), Blue (B), Hue (H), Saturation (S), Value (V), RGB (Red, Green, Blue), HSV (Hue, saturation, Value), HIS (Hue, Saturation, Intensity). Generally, to identify the disease from a leaf, two tasks are required. First, image processing task and second machine learning task. At each step of image processing task, different techniques are available. Through study of different literatures, we came to know that no dataset for this problem is available freely. Most authors [21],[5],[6],[11],[22] have mentioned in their works that they had captured images for their experiments from rice fields. Some of the authors, e.g., in [23],[24], used some standard database such as IRRI (International Rice Research Institute) and another dataset available at shutterstock.com was used by [25].

TABLE I. Analysis of Image Processing Operations applied in Rice Disease Identification

No	Ref	Dataset	Diseases	Preprocessing	Segmentation	Edge Detection	Feature Extraction	Background of images	Color space
1	[21]	50 images of each disease	RB, SB, BS	Yes Digitization, Quantization	Yes Thresholding	Yes Sobel	Yes Differences of two color space,area,roundness Shape complexity,Length concavity,longer axis,shorter axis	Dark	RGB
2	[23]	Not specified	RB	Yes Weiner filter, Contrast enhancement using histogram equalization	Yes K-means	No	Yes Mean, standard deviation	Not specified	RGB
3	[5]	72 imagesfor each disease	BLB SB RB	Yes Resolution reduction, remove noise using median filter	Yes Otsu's method	No	Yes Area,peimeter,contrast, uniformity,entropy,inverse difference,linearity correlation,rectangularity, compactness,elongation ,roundness	Non-uniform natural background	RGB
4	[25]	6 images of each disease	BS, SR, RB	No	Yes Fermi energy based	No	Yes R,G,B mean and standard deviation of infected and background pixels	Not specified	RGB
5	[6]	500 images of each disease	RB, BS	Yes mean filter for image enhancement	Yes Otsu's method	No	Yes Radial distribution of hue from center to boundary of spot	Not specified	RGB, Grayscale
6	[26]	60 samples	BS, NBS, BLB, RB	Yes Green plane extraction, median filtering, binarization	Yes 8-connected component Labeling	No	Yes Weight matrix from PCA, mean of H,S,V	White BG	RGB HSV
7	[24]	BLB-409, RB-360 BS-319 Healthy-239	BLB, RB, BS	Yes Crop area of interest, resize image(50*50), Smoothing	No	Yes Canny	Yes Color, shape	Non-uniform natural bg	RGB CIE LAB

8	[10]	20 images	BS	No	No	No	Yes R,G,L component	Not specified	RGB, LAB
9	[27]	BLB-55 BS-37 RB-42	BLB, BS, RB	Yes Image enhancement,otsu Method	Yes Thresholding, masking	No	Yes Fraction covered by disease on leaf, mean,std of R,G,B,Mean of H, S,V	Black	RGB, HSV
10	[11]	RB-14 BS-37 NBS-47	RB, BS, NBS	Yes Median & morphological operators	Yes Otsu's method	No	Yes No. of spots, width & length of spot, color difference between two color space RGB & LAB	Not specified	RGB LAB
11	[22]	Not specified	BS, LB, Tungro, RB	Yes Cropping, converting to gray scale, enhancement using laplacian	No	No	Yes Fuzzy entropy	Not specified	RGB
12	[28]	Not specified	BS, RB	No	Yes Entropy based bi-level thresholding,boundary detection, spot detection, zooming algorithm	No	No	Non uniform Natural	HSI
13	[29]	500 images	BS, RB, SR, BLB	No	Yes Fermi energy based	No	Yes Mean,standard deviation, position of infection changes of pixel values of spot w.r.t pixels, central moment	White	RGB

Many diseases can occur on rice plants, but mostly Rice Blast [21],[5],[25],[23],[6], Brown Spot [21],[25],[6],[26] and Bacterial Leaf Blight[5],[26],[24],[27] are frequently occurring on the plants.

After getting the images, image preprocessing step is required, as discussed earlier. Researchers have used median filter to remove or weaken the noises from the images[5],[26],[11]. Median filter uses a mask of 3X3 window which runs over the whole image and replaces all the pixels with median of the neighbourhood pixels. In some cases, while applying median filter, if spots on the leaf are very small, then they will get removed or blurred out from the leaf. Authors have done image enhancement using different techniques such as mean filtering [6], histogram equalization [23], laplacian filter [22]. In [26], the authors used green plane extraction. They were trying to enhance the affected portion of leaf by extracting the green part of the image because greenness is more affected while infection occurs on leaf. In [11], for removing unnecessary spots median filter and morphological operators are used.

Researchers used different techniques of segmentation as per their requirements. Table II presents comparison of various segmentation techniques used in detection of leaf disease. Generally, authors used thresholding technique to segment the diseased portion from the leaf, e.g., in [21],[27],[28]. Thresholding technique uses a specific threshold to bifurcate the image into two regions. In thresholding, specifically Otsu's thresholding is useful because Otsu's thresholding automatically sets the threshold value. In [5],[11],[6] Otsu's thresholding was used. The drawback of Otsu's method is that it is subjective to image content. That means the threshold that works on some images accurately may not work on other images due to presence of different grey levels in the image. K-means clustering is used for the image segmentation also, e.g., in [23]. K-means clustering groups similar color pixels into one cluster, while groups different color pixels into different clusters. K-means clustering can provide accurate result, for leaf images, as compared to other segmentation techniques due to its ability to cluster based on distance between two pixels. As, generally, disease portion and green portion have different color intensities, K-means clustering can generate appropriate clusters.

In [21], Sobel edge detection is used to detect edges of leaf. These edge pixels are used in calculating the density of the image. Density is calculated by divide the number of edge pixel with area of the image. In [24], they carried out experiment on different edge detection algorithms such as Roberts, Canny , Sobel and prewitt. After that they concluded that canny edge detector gives good result compared to other edge detectors, because it could detect more detail information of plant leaf.

After extracting disease portion, the features from the disease portion are extracted. Generally, mean and standard deviation of R, G, and B components of diseased portion is mostly used as features [25],[23],[29]. The mean of each component may vary from disease to disease, because the color of disease portion is different for different diseases. For example, brown spot has brown color intensity, and that disease has different ranges of mean and standard deviation as compared to other disease like leaf blight, which has yellowish type color. The shape of disease portion can also vary from disease to disease. The area, e.g., used in [21],[5], of the disease spots are calculated from the binary image of disease spots. Texture features such as contrast, uniformity, and linear correlation are also important features [5].

Generally, the image with uniform background gives good result and has less computational time as compared to non-uniform background. Mostly researchers used uniform background in their research work, e.g., in [21],[26],[29].

TABLE II. COMPARATIVE STUDY OF SEGMENTATION TECHNIQUES

	Technique name	Thresholding Type	Segmentation type	Complexity	Segmentation effect	Merit	Demerit
1	Otsu's method	Global	Thresholding	Very High	Good/stable	Regardless of uniformity & shape measures, it works on real world images	Takes more time in processing
2	Fermi energy Based	Global	Thresholding	Low	Better compared to Otsu and k-means	Overcomes the limitation of selecting proper threshold value	Only works when non-uniform illumination is present
3	K-means	Local	Clustering	Low	Accurately distinguish infected & uninfected regions of plants	Minimizes sum of square distance between object and centroid	Difficult to predict k with fixed number of clusters.
4	Grey-level thresholding	Global	Thresholding	Normal	More accurate compared to Otsu's method	grey level transformation(2G-R-B) provides contrast for disease region and background.	Every time needs to select proper threshold value for getting better result in segmentation.
5	Fuzzy c-means	Local	Clustering	High	Better compared to Otsu and k-means	Uses partial membership; therefore, more useful for real problems.	Sensitive to initialization condition of cluster number and cluster center

VI. SURVEY AND ANALYSIS OF MACHINE LEARNING OPERATIONS APPLIED FOR RICE DISEASE IDENTIFICATION

This section presents survey of machine learning techniques used in classification of rice plant diseases. The survey is carried out with the following criteria: types of classifier, parameters for classification, input to classifier, and accuracy of classifier. Table III presents the summary of various machine learning techniques applied for disease classification. We use following notations in the table: Support Vector Machine(SVM), Ensemble Learning(EL), Quadratic Discriminant analysis(QDA).

Researchers used different classifiers as per their requirement. Generally, authors use support vector machine for classification of diseases, e.g. in [5],[6],[23],[30]. Support vector machine is well suitable for high dimensional data. Support vector machine is strong enough in case of data has some distortion[16]. Support vector machine has different kernel functions. For two class classification linear function is most suitable kernel function, while for multi class classification radial basis kernel function is more preferable kernel function. While some authors used neural network for classification, e.g. in [10],[22],[24],[27]. Neural network is used in case of large number of training samples. Neural network requires more processing time[16]. Rule generation method is least used classifier. Generally, rule generation classifier is based on the production rules. We need to create some production rules in this method. For example, if color and shape features of trained dataset image do match with the test image then, that test image falls under that specific disease category. Input to the classifier is generally image features which is extracted from the diseased portion, but in some cases like in [10],[28] authors give input as image pixels.

TABLE III. ANALYSIS OF MACHINE LEARNING OPERATIONS APPLIED IN RICE DISEASE IDENTIFICATION

	Ref.	Types of classifier	Parameters for classification	Inputs	Accuracy
1	[21]	Nearest Neighbour	RGB range, shape, length, width, diameter	Membership Function	Rice Blast-80% Rice sheath blight-60% Brown spot-85%
2	[5]	Support Vector Machine	Radial Basis Kernel Function	Image Features	Model1-97.2%, Model2-88% Model3-11.1%
3	[30]	Support Vector Machine, Neural Network, Ensemble Learning, Quadratic Discriminant analysis	Default Parameters	Image Features	SVM, EL, QDA-85% NN-80%
4	[23]	Support Vector Machine	Not Specified	Image Features	82%
5	[25]	IF-Then Classifier	Color and shape features	Image features	75%
6	[6]	Bayes Classifier Support Vector Machine	Not Specified	Image Feature	Support Vector Machine-68.1% Bayes classifier -79.5%
7	[10]	Backpropagation Neural Network	3 hidden layers	R, G and L pixels	90%
8	[27]	Backpropagation Neural network	Minimum require gradient- 9.7909e-13 Epoch-87	Image Features	100%
9	[31]	Production Rule with forward	Not specified	Image Features	Local Entropy-100%

		chaining			Otsu-39.9%
10	[29]	Rule Generation	(Feature, value) pair	Image features	92.29%
11	[28]	Self-organizing map neural network	50 epochs	Gray values of pixels	RGB spot -92% Fourier transform of spot-84% rotation of 50 %spot- 82%
12	[22]	Probabilistic neural network	Radial basis function for pattern layer	Image Features	Brown spot-100%, Leaf blast-100% Tungro-76%, Bacterial leaf blight-96%
13	[24]	Backpropagation neural network	Input values -Binary,Input Neurons-2500 Hidden Neurons-500,Output Neurons-3 Learning Rate -0.2,Activation Function -Sigmoid,MSE < 0.0001	Image Features	74.21%

VII. PROPOSED WORK

This section presents the proposed work for the rice disease identification. In our proposed system, we intend to detect three rice diseases namely brown spot, bacterial leaf blight, and leaf smut. The block diagram of the proposed work is shown in Fig.3. We discuss the processing steps of our proposed work in following subsections. We also point out the reason for taking a particular decision.

A. Image Acquisition

We have collected samples of rice plant, both normal and having diseases, from a village called Shertha near Gandhinagar, Gujarat, India. We could get desired leaves of rice plantsfrom the rice field at the start of the winter season in India, in November 2015. We used NIKON D90 Digital SLR camera with 12.3 effective megapixels. We captured images of the leaves on white background, under direct sunlight. We have prepared total 145 images in our database containing 30 images of healthy leaf, 46 images of bacterial leaf blight, 44 images of leaf smut, and 25 images of brown spot. As November was harvesting period for rice crops in Gujarat, we could get only a few hundred leaves.

B. Image Pre-processing

As shown in Fig. 3, the RGB image is converted into HSV color model. Then, we choose S component of the HSV image because S component does not contain the whiteness. After that, we create a mask such that the mask removes and makes all the background pixels as zeros.

C. Image Segmentation

K-means clustering is used for the segmentation. Three clusters are expected from the leaf image, (1) background, (2) infected, and (3) green clusters. We apply K-means clustering on the Hue component of the HSV image. For some images, normal K-means algorithm could not produce desired clusters, three segments. To produce accurate segments, we feed the centroid value of each desired cluster, which we find based on histogram analysis of Hue components of the leaf image.

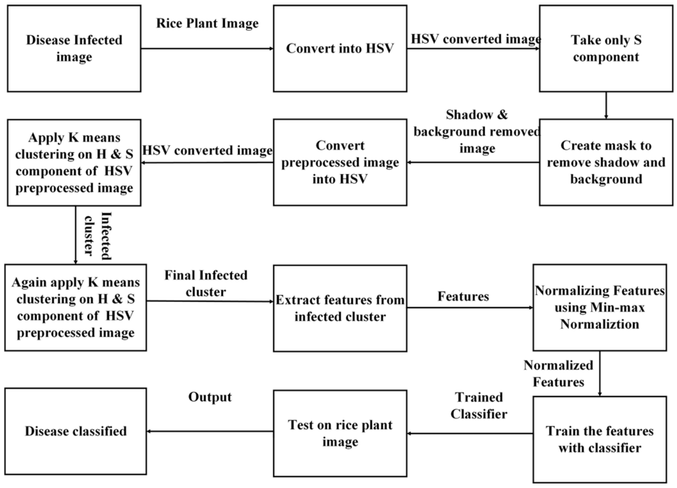

Fig. 3. Block Diagram of the Proposed Work

D. Feature Extraction

Features have crucial role in differentiating one disease from another. We intend to use color features such as mean and standard deviation. We also plan to use texture features such as Kurtosis, skewness, cluster prominence, and cluster shade. For extracting texture features, gray level co-occurrence matrix will be used. Another features that we are empirically exploring are the number of diseased spots and the number of pixels of disease portion falling under 7 pre-defined color ranges, which we decide based on manual color analysis of all the types of diseases that we want to classify. At present, we have extracted around 50 features from few images, and we are validating our implementation, done in Matlab, of feature extraction step.

E. Classification

We will use Support vector machine for the classification. Support vector machine is supervised learning approach. It classifies the training data based on the classes given as training class labels. Linearly separable classes can be identified using a hyper plane while for the data points which are not linearly separable can be handled using kernel function. In our work we have three classes(diseases). We use Gaussian kernel for multiclass classification.

VIII. CONCLUSION

Rice plant diseases can incur tremendous amount of loss in agriculture if enough attention is not given. Using computer and communication technologies, an automated system can be built which can provide early notification of disease. In the same direction, we tried to provide our contributions in image processing and machine learning aspects of such system. We have studied that various alternatives exist for various operations in image processing and in machine learning. This paper reviewed and summarized techniques of the image processing and machine learning that have been used in disease identification. We found that extraction of disease region from the leaf image is the driving step, for which we have studied and compare various segmentation techniques. We utilized our survey and study, presented in this paper, to propose our work in the same direction. The paper presented detailed schematic diagram of the proposed work and discussed important steps. At present, we are working on completing the implementation of the proposed work. A combination of image processing and machine learning techniques can give opportunities to researchers to address problems in various domains that affect to society directly or indirectly.

REFERENCES

[1] "Agriculture Economics and Importance of Agriculture in National Economy."[Online].Available:http://agriinfo.in/?page=topic&superid=9&topicid=185. [Accessed: 23-Nov-2015].

[2] "Agriculture Sector in India." [Online]. Available: http://www.ibef.org/industry/agriculture-india.aspx. [Accessed: 23-Nov-2015].

[3] L. P. Gianessi, "Importance of Pesticides for Growing Rice in South and South East Asia," pp. 30–33, 2014.

[4] "Rice Production (Peace Corps): Chapter 14 - Diseases of rice." [Online]. Available: http://www.nzdl.org. [Accessed: 23-Nov-2015].

[5] Q. Yao, Z. Guan, and Y. Zhou, "Application of support vector machine for detecting rice diseases using shape and color texture features," in International Conference on Engineering Computation, pp. 79-83 2009.

[6] 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, pp. 460-463, May 2012.

[7] D. Groth and L. S. U. Agcenter, "Rice Disease Identification Photo Link."

[8] "IRRI - Rice science for a better world." [Online]. Available: http://irri.org/. [Accessed: 23-Nov-2015].

[9] A Fakhri, A Nasir, M Nordin, A Rahman and A R.Mamat, "A study of image processing in agriculture application under high performance computing environment," International Journal of Computer Science and Telecommunications, vol. 3, no. 8, pp. 16-24, Augest 2012.

[10] L. Liu and G. Zhou, "Extraction of Rice Leaf Disease Image Based on BP Neural Network," in Computational Intelligence and Software Engineering, Wuhan, 2009.

[11] N. N. Kurniawati, S. Norul, H. Sheikh and S. Abdullah, "Investigation on image processing techniques for diagnosing paddy diseases," in Soft Computing and Pattern Recognition, 2009.

[12] P.L.sahu, A. Singh and K.L. Sinha, "a survey on data mining techniques for classification," vol. 2, no. 1, pp. 65-70, 2015.

[13] J. Han,M. Kamber, and J. Pei. Data mining: concepts and techniques: concepts and techniques. Elsevier, 2011.

[14] K.I. Rahmani, N. Pal, and K. Arora. "Clustering of Image Data Using K-Means and Fuzzy K-Means." International Journal of Advanced Computer Science and Applications 5, no. 7 (2014): 160-163.

[15] J. Garci and J. A. Barbedo, "Digital image processing techniques for detecting, quantifying and classifying plant diseases," Spinger Plus, vol. 2, pp. 1-12, 2013.

[16] R. Pandey, S Naik, and R.Marfatia. "Image Processing and Machine Learning for Automated Fruit Grading System: A Technical Review"." International Journal of Computer Applications 81, no. 16 (2013).

[17] A.Vibhute, and S. K. Bodhe. "Applications of image processing in agriculture: a survey." International Journal of Computer Applications 52, no. 2 (2012).

[18] S.D.Khirade and A.B.Patil , 2015, February. Plant Disease Detection Using Image Processing. In International Conference on Computing Communication Control and Automation (ICCUBEA), 2015 (pp. 768-771). IEEE.

[19] H. Wang, G. Li, Z. Ma, and X. Li. "Image recognition of plant diseases based on backpropagation networks." In 5th International Congress on, Image and Signal Processing (CISP), 2012pp. 894-900. IEEE, 2012.

[20] H. Al-Hiary, S. Bani-Ahmad, M. Reyalat, M. Braik and Z. alrahamneh. "Fast and accurate detection and classification of plant diseases." Machine learning 14 (2011).

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

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

[23] A. K. Singh, A. Rubiya, B. S. Raja,,"classification of rice disease using digital image processing and svm," International Journal of Electrical and Electronics Engineers, vol. 07, no. 01, 2015.

[24] M.A.Kahar, S. mutalib, and A.Rahman. "Early Detection and Classification of Paddy Diseases with Neural Networks and Fuzzy Logic." pp. 248–257.

[25] C.K.Charliepaul, "classification of rice plant leaf using feature matching," vol. 1, pp. 290-295, 2014.

[26] T. Suman and T. Dhruvakumar, "classification of paddy leaf diseases using shape and color features," vol. 07, no. 01, pp. 239–250, 2015.

[27] J.W Orillo, J.D. Cruz, L. Agapito, P J Satimbre, and I. Valenzuela. "Identification of diseases in rice plant (oryza sativa) using back propagation Artificial Neural Network." In Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management (HNICEM), 2014 International Conference on, pp. 1-6. IEEE, 2014.

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

[29] S.Phadikar,J. Sil and A. K. Das. "Rice diseases classification using feature selection and rule generation techniques."Computers and electronics in agriculture 90 (2013): 76-85.

[30] G. Maharjan, T. Takahashi, and S. H. Zhang. "Classification methods based on pattern discrimination models for web-based diagnosis of rice diseases."Journal of Agricultural Science and Technology 1, no. 1 (2011): 48-56.

[31] N. N Kurniawati, and S.Abdullah. "Texture analysis for diagnosing paddy disease." In International Conference on Electrical Engineering and Informatics, 2009. ICEEI'09., vol. 1, pp. 23-27. IEEE, 2009.