

Paddy Disease Analysis Using Image Processing Technique

D.Rechel, M.Phil¹, Dr .A.Shaik Abdul Kadhir, M.Sc.,
M.Phil, Ph.D²

Research Scholar, Khadir Mohideen College, Adirampattinam.

Head & Associate Professor Of Computer Science, Khadir Mohideen College, Adirampattinam.

Submitted: 15-08-2021

Revised: 29-08-2021

Accepted: 31-08-2021

ABSTRACT

Paddy is a monocotyledonous angiosperm and it belongs to *Oryza*. It contains more than 20 species, yet just two species are alluded to as development. In this proposed work different data set is considered with kNN classifier, in this experimentation, the features extracted from red component is used for training and the features extracted from green component is consider for testing which has given a success rate of 87%. In agricultural field recognition of healthy or diseased plant was the major constrain. The control period was habitually missed because the traditional approaches are taking long time to analyse. Because an expert has to visit agricultural field repeatedly, it may take more time to analyse,

and it may be burden to the farmer. The Proposed system we analyze the paddy disease .

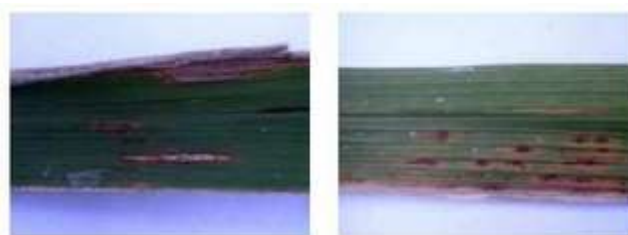
Keyword: Paddy Disease, KNN Clasifier, Feature Extraction.

I. INTRODUCTION

PADDY DISEASE:

Leaf Blast (*Pyricularia grisea*)

Leaf blast (LB) caused by *Pyricularia grisea*, a fungal pathogen. A blackish, rhombus-shaped, brown spot with a white colored central patches will appears at the leaves of the sick plan. Leaf blast is difficult to control because it's ability to forms pathogen strains quickly .



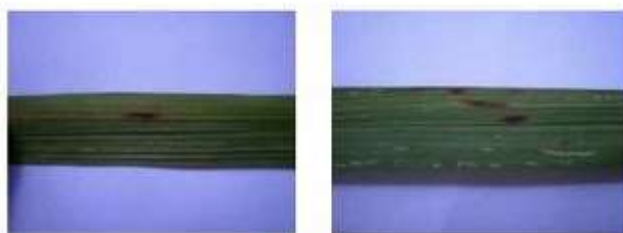
Leaf Blast

Fig 1.4.1 leaf blast

Brown Spot (*Helminthosporium oryzae*)

Brown spot (BS) caused by the fungus *Helminthosporium oryzae*. This disease can cause death in young rice plants and reduces grain

quality. Leaves of the rice plants suffering from this disease will have an oval shaped or round shaped brown colored patches about the dimensions of sesame seeds.



Brown Spot

Fig 1.4.2 Brown spot

Bacterial Leaf Blight (*Xanthomonas campestris*)



Bacterial Leaf Blight

Fig 1.4.3 bacterial leaf

Bacterial blight (BLB) may be a disease caused by the bacterium *Xanthomonas campestris*. This disease is widespread and lowering the yield

to 36%. HBD disease produces two characteristic symptoms: the crackle and blight. Crackle may be a symptom that happens in old plants

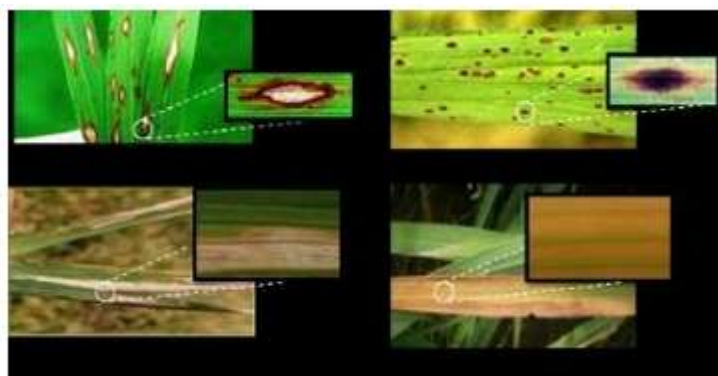
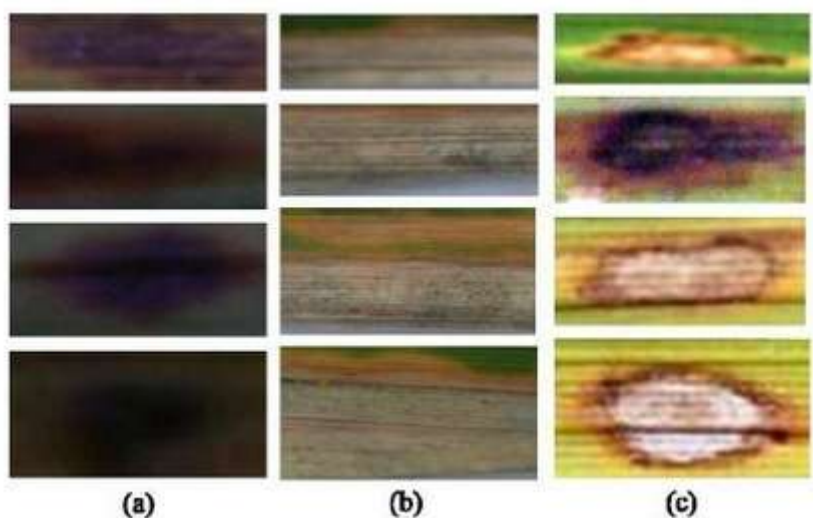


Fig 1.4.4 paddy disease

Tungro (TG) is one of the important diseases in rice. It is very destructive. It caused by tungro bacilliform virus (RTBV). Depends on the phase of the infected plants, tungro can cause 5-70% yield loss. The younger the plants is, the loss caused by tungro are going to be greater. Tungro attacks prominent symptom is discoloration of leaves and stunted plant growth. Diseased plant leaf color varies from slightly yellow to orange. The typical symptoms are determined by the extent of resistance of sorts, environmental conditions, and therefore the growth stage when the plants infected. From above description, we already seen

the overall characteristic of every disease. In practice, diagnosing a disease visually involves more than one lesion characteristic i.e. the location of the symptoms, the spots or lesions shape, the dimensions of the spots or lesions, and therefore the color of the spots or lesions. In this research, we only focused on describing the colour texture of the disease lesions. This, however, doesn't mean that color texture is enough to differentiate the diseases. Some diseases may have an identical color and may only be differentiated by using the form. The most obvious example of this kind of disease are brown spot.

No	Pest	Area of Attack (ha)
1	Rice Stem Borers	57,875
2	Brown Planthopper	11,351
3	Rats	54,300
4	Tungro	4,994
5	Leaf Blast	31,383
6	Bacterial leaf blight	43,719



II. LITERATURE REVIEW

2.1 IDENTIFICATION OF HEALTHY AND DISEASED PADDY LEAVES USING Knn CLASSIFIER

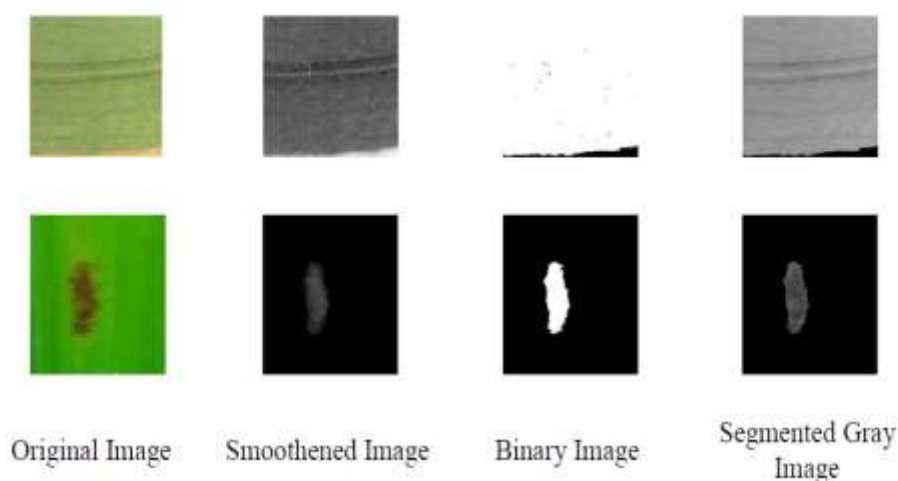
AUTHOR: Shreekanth K N

Rice is a monocotyledonous angiosperm and it belongs to *Oryza*. It contains more than 20 species, yet just two species are alluded to as development ie *Oryzasativa* and *Oryzaglaberrima* developed in south-east Asia and West African nations. Basically rice was developed in the tropical locale of Asia, and from the past record rice is from 5000 years BC, and there after reached out to calm areas . Rice is a fundamental staple nutrition food in Asia. In an Asian country rice is the major leading food which is grown all over the world ie 90%. Three billion Asian people were intake caloric in the ratio of 35 – 60% . Plant development was suggestively decreased with compared to un-inoculated plants and inoculated plants. Blast disease indications of Rice are also produced in inoculated plants . Rice Brown Spot is the evolving disease in low fertility and marginal areas which causes substantial yield loss . Initially brown spot lesions and blast diseased lesions are difficult to distinguish. Blast lesions turn to

diamond or spindle shape after maturity .Rice (*Oryzasativa*) blast, instigated by *Magnaportheigrisea*, which is one of the most harmful agent of rice. Many studies have been conducted to resist mechanisms of rice to the blast fungus and thus several antifungal substances have been isolated from rice leaves . In particular, seven parasites *Curvularialunata*, *Magnaportheoryzae*, *Fusariummoniliforme*, *Helminthosporiumoryzae*, *Nigrosporaoryzae*, *Aspergillusniger* and *Alternariaalternata* were disconnected from influenced leaves and seeds contaminated leaves of five rice assortments viz., IRRI-8, IRRI-6, KS-282, DR-82, and Shua. N. *Oryzae* is disconnected just from IRRI-8 and IRRI-6 rice assortments. Among these growths, M. *Oryzae* was, for the most part, disengaged both from leaves and seeds of the rice assortments considered. M. *Oryzae* was detected with maximum frequency from seeds and leaves of IRRI-6 .Brown Spot diseases is one of the most significant diseases in rice, Brown spot is caused by the fungus *Bipolarisoryzae* (earlier identified as *Helminthosporiumoryzae*)can affect all plant parts of the paddy which will affect the quality of grain and yield losses. Brown spot symptoms primarily

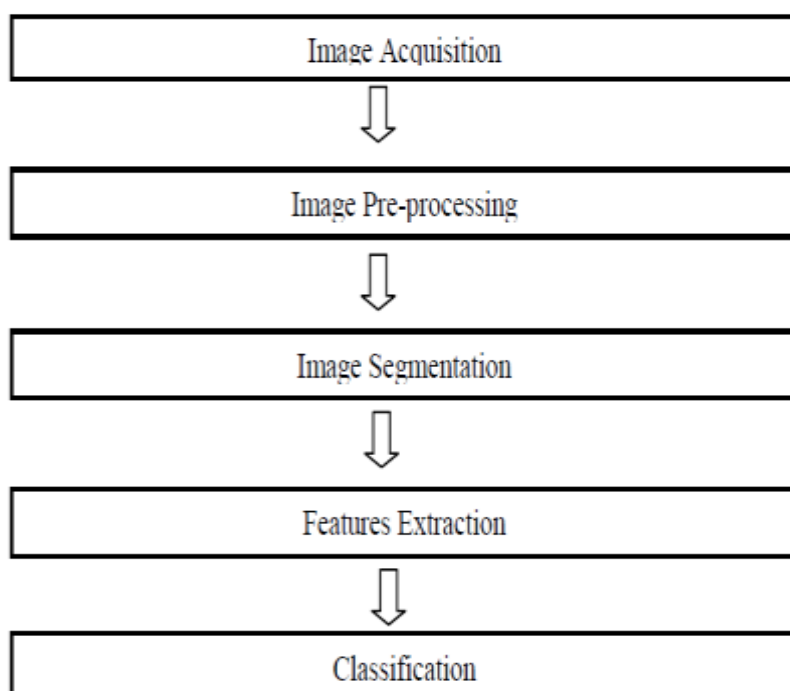
appear as small circular to oval spots on the leaves. Yield loss, shape, size, and color these depend on and and these are estimated based on environmental conditions and rice variety, 16 to 40% of yield loss observed in Florida. Large spots have a light grey centre enclosed by a dark to reddish-brown margin or reddish-brown while small spots are dark brown to reddish brown. Elder spots may have a bright yellow halo surrounding the lesion. Spots on the hulls and leaf sheath are similar to those on the leaves. Diseases are most important causes for quality and quantity reduction in agricultural products. Plant diseases also can harm and wreck the agricultural fields. The tungro

is the most important rice virus disease; it is transmitting through green leaf hopper and it is the one of most destructive diseases of rice at the early stage of the tungro. Disease symptoms are absent and symptoms appear only in the later stages . Rice plants infected with tungro symptoms including stunting growth which leads to yellow to orange discoloration and reduced tillering. Leaves may discolour and shows irregular shaped dark-brown blotches. The younger leaves how striping or mottling along with intervarnalchlorosis. Bacterial blight is one of the most devastating diseases of rice worldwide and is found both in tropical and temperate regions .



The disease is accountable for a loss of 20-30% in different countries. In India losses in yield varied from 6-60% in different states depending upon stage and severity of infection and type of cultivars . Leaf blight is one of the most common syndrome, it begins like water-soaked stripes on the leaf blades. Then stripes enlarged in area and become yellow to white and enclosed entire leaf blade. Water-soaked margins are small circular lesions, they appear in severe infections, and infected plants produce poor quality of grain. Bacterial leaf blight is produced by *Xanthomonas oryzae*. It is Gram-negative, non-spore forming rod .Bacterial Brown Stripe also

known as bacterial stripe occurs in nurseries and also distributed widely in rice growing countries .The symptoms of this disease in seedling stage are inhibition of germination occurring brown stripes interveinally or margins of leaf sheath is curving and abnormal elongation of mesocotyl. Manifestations were covered after the contaminated seedlings were transplanted to paddy fields. The regular event of the illness was seen where rice plants were submerged by overflowed water . Bacterial dark colored stripe is caused by *Pseudomonas syringae*pv. *panici*. is a Gram-negative, non-spore-shaping, non-embodied pole .



2.2 Paddy Diseases Identification with Texture Analysis using Fractal Descriptors Based on Fourier Spectrum

AUTHOR: Auzi Asfarian, Yeni Herdiyeni

The efforts to increasing the quantity and quality of rice production are obstructed by the paddy disease. This research attempted to identify the four major paddy diseases in Indonesia (leaf blast, brown spot, bacterial leaf blight, and tungro) using fractal descriptors to analyze the texture of the lesions. The lesion images were extracted manually. The descriptors of 'S' component of each lesion images then used in classification process using probabilistic neural networks. These techniques achieved at least 83.00% accuracy when identifying the diseases. This method has a potential to be used as one of the feature if it combined with other features, especially when two diseases with relatively same color involved.

The efforts to increase the quantity and quality of rice production to satisfy the increasing needs of rice in Indonesia experienced several obstacles, one of which is the attack of the diseases on paddy fields. Indonesian Directorate General of Food Crops stated that during the period of October 2011 to March 2012, 80,096 hectares of paddy fields exposed to attack by three major paddy diseases in Indonesia: tungro, leaf blast, and leaf blight. To control these diseases and to minimize

the impacts of the attacks, the diseases must be identified quickly. Unfortunately, experts who are able to identify the diseases are often unavailable in some region. Computer vision is a potential solution to tackle 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. This method also nondestructive and the results are consistent. This method involves the extraction the features of the said disease lesion. The common paddy lesion features are the texture, the color, the position, or the size of spots or lesions. Some research combined more than one of these features. For example, used the texture, color, and shape to recognize blast, sheath blight, and brown spot, the three major rice diseases in Sri Lanka, with 70% accuracy. used a color features (e.g. boundary color and spot color) to recognize blast, brown spot, and narrow brown spot diseases and achieved 87.5% accuracy. proposed a new technique to analyze the texture using fractal descriptors based on image Fourier spectrum. When tested to four different datasets (Brodatz, USPTex, OuTex, and plant leaves), this method is more accurate and faster than any other fractal descriptor estimation techniques. This research

attempted to identify the four major paddy diseases in Indonesia using fractal descriptors proposed by and assess the performance of said method. The four diseases are leaf blast.

III. PROBLEM DESCRIPTION EXISTING SYSTEM

In agriculture sector where farmers and agribusinesses need to make innumerable decisions a day and complex complexities involves the varied factors influencing them. An essential issue for agricultural planning intention is that the accurate yield estimation for the various crops involved within the planning. Data mining techniques are necessary approach for accomplishing practical and effective solutions for this problem. Agriculture has been a clear target for giant data. Environmental conditions, variability in soil, input levels, combinations and commodity prices have made it all the more relevant for farmers to use information and obtain help to form critical farming decisions. The texture features utilized in this work are, LBP, Haar Wavelets, GLCM and Gabor features. The attributes utilized in these texture features are

explained within the subsequent sections. Results obtained from each texture feature set and best discriminative subset of features from HSI and YCbCr color models are selected empirically supported combination of features. The below mentioned algorithm discuss about proposed methodology which covers segmentation, feature extraction and classification.

Algorithm

1. Transform a RGB arecanut image to HSI and YCbCr color model.
2. Extract Saturation component from HSI for segmentation of arecanut from background.
3. Threshold based segmentation is used based on Global image threshold using Otsu method.
4. The color components extracted from HSI and YCbCr are Hue, Saturation, Intensity, Yellow, Chromatic Blue and Chromatic red.
5. Enumerate all these color components. Segmented image is multiplied with color component in iteration.
6. From each color component Haar Wavelets, GLCM, Gabor texture features are extracted.

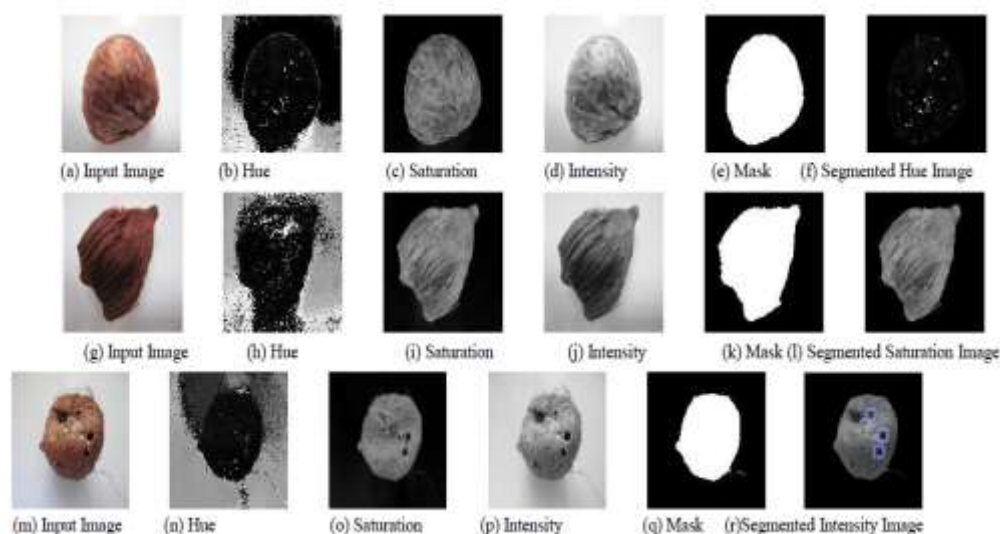


Fig 3.1 feature extraction

Wavelets Features used in this work

Mean Value of approximation, horizontal, vertical and diagonal energy	$\frac{1}{mn} \sum_{i=1}^m \sum_{j=1}^n x_{ij}$
Mean of Decomposition Vector	$\frac{1}{n} \sum_{i=1}^n x_i$
Variance of Decomposition Vector	$\frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2$
Entropy of Decomposition Vector	$-\sum_{i=1}^n p(x_i) \log_2 p(x_i)$

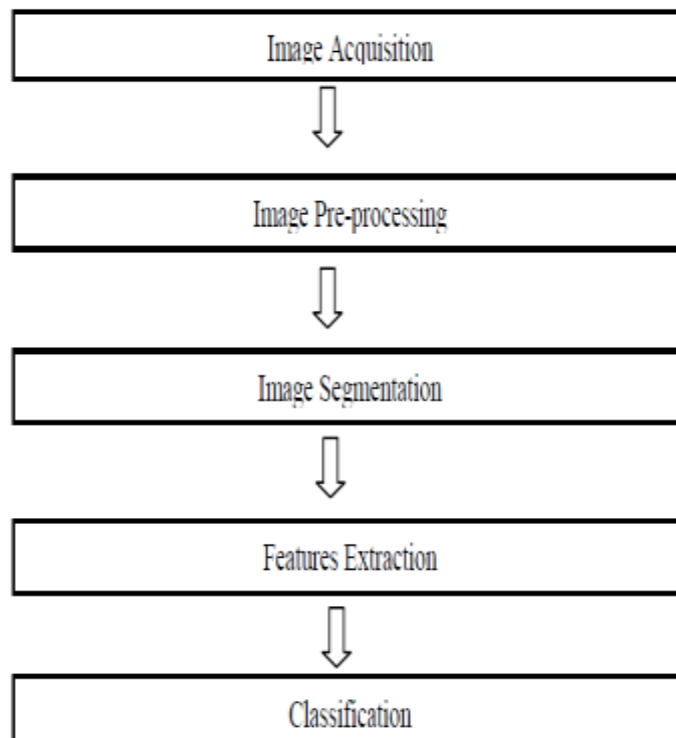


Fig 3.1.2 Architecture

A. Image acquisition:

Sampled Images are captured using sony digital camera with 18.1 megapixels in day light

because clear view of leaf object is required for analysis. Sample diseased and healthy paddy leaves are shown in figure 2.

B. Image Pre-processing:

Pre-processing is the preliminary step before segmentation and feature extraction. In this proposed work images are resized to 255 X 255 to improve the accuracy of an algorithm. Individual RGB components are extracted for image smoothening. Here the G and B components are interpolated with scalar values and the resultant value is subtracted with R component and multiplied with a value shown in equation (2). Images are smoothened and then enhanced. In resultant image the healthy part and diseased part can be easily identified because the diseased part appears darker compared to other part of the leaf using equation (2). Same type of other calculation for R, G, B components are conducted .

$$SIM = 2 X (R - (0.75 X G - 0.25 X B)) (2)$$

C. Image segmentation:

In image analysis and pattern recognition segmentation is the major task, partitioning of an image into groups of pixels based on similarity measurement is known as segmentation. It helps in the feature extraction from the segmented image based on the interested area or based on the objective which is considered for the experiment . Pre- processing of images is significant task. In proposed methodology image smoothening were carried out using equation (2). For resultant images segmentation are carried out by using global thresholding method of Otsu segmentation. Experimental segmented images .

<i>Normalized Block Size</i>	<i>Sample data set</i>	<i>Accuracy</i>
255 X 255	Gray images	77.71%
	Similarity difference of Gray with Red component	70.48%
	Similarity difference of Gray with Green component	75.30%
	Similarity difference of Gray with Blue component	78.91%
	Red component	79.57%
	Green component	79.51%
	Blue component	76.50%
32 X 32	Gray images	84.93%

D. Feature extraction:

The local binary pattern texture feature information extraction was proposed by Ojala et al.,.Based on theory, texture has locally two corresponding parts like a pattern and its strength with the goal of texture

- Image Acquisition
- Image Pre-processing
- Image Segmentation
- Features Extraction

classification and then it is extended to face detection, face recognition, face expression recognition, disease identification etc.

The advantages of LBP are it is invariance to monotonic gray-scale modifications, convenient multi-scale extension and minimum computational. Traditional and statistical structural methods are well designed in LBP. In an input image each processing pixel is matched with its eight neighbors and ones whose intensities exceed the processing pixels are marked as 1, else as 0, in resultant a circular point features consisting of only binary bits were measured. Naturally the feature ring is considered as row vector, and binomial weight assigned to each bit. Row vector is transformed to decimal for further work. Local Binary Pattern using circular neighbourhoods and linearly

incorporating the pixel values allows the choice of any radius, R , and number of pixel in the neighbourhood, P , to form an operator, which can model large scale structure. Basic LBP is noise sensitive, because the threshold value is the processing pixel, so average of pixels including

processing pixel that encompass by an LBP operator as a threshold. The mathematical model for LBP is shown in equation (3)

Where gc is the average gray value of the pixels and gp is the intensity value of pixels in eight neighborhood.

$$\begin{aligned}d(GRAY, R) &= GRAY(x, y) - R(x, y) \\d(GRAY, G) &= GRAY(x, y) - G(x, y) \\d(GRAY, B) &= GRAY(x, y) - B(x, y)\end{aligned}$$

Where GRAY – Gray component,
R – Red component,
G – Green component,
B – Blue component.

3.2 PROPOSED SYSTEM:

Classification algorithm in R mining

Data Mining may be a technique utilized in various domains to offer aiming to the available data Classification may be a data processing (machine learning) technique wont to predict group membership for data instances. In this paper, we present the essential classification techniques. Several major sorts of classification method including decision tree, Bayesian networks, k-nearest neighbour classifier, Neural Network, Support vector machine. The goal of this paper is to provide a review of different classification techniques in data mining. Keywords— data processing , classification, Support vector machine (SVM), K-nearest neighbour (KNN), Decision Tree.

3.2.1 K-NEAREST NEIGHBORS ALGORITHM

Known as: Ibk algorithm, Nearest neighbors classifier, K-NN

In pattern recognition, the k-Nearest Neighbors algorithm (or k-NN for short) may be a non-parametric method used for classification and regression. In both cases, the input consists of the k closest training examples within the feature space. The output depends on whether k-NN is employed for classification or regression: k-NN may be a sort of instance-based learning, or lazy learning, where the function is merely approximated locally and all computation is deferred until classification. The k-NN algorithm is among the only of all machine learning algorithms.

3.3 PROPOSED TOOL

INTRODUCTION TO R

R may be a language and environment for statistical computing and graphics. It is a GNU project which is similar to the S language and environment which was developed at Bell Laboratories (formerly AT&T, now Lucent Technologies) by John Chambers and colleagues. R are often considered as a special implementation of S. There are some important differences, but much code written for S runs unaltered under R.

R provides a good sort of statistical (linear and nonlinear modelling, classical statistical tests, time-series analysis, classification, clustering, ...) and graphical techniques, and is very extensible. The S language is usually the vehicle of choice for research in statistical methodology, and R provides an Open Source route to participation therein activity.

One of R's strengths is that the ease with which well-designed publication-quality plots are often produced, including mathematical symbols and formulae where needed. Great care has been taken over the defaults for the minor design choices in graphics, but the user retains full control.

R is out there as Free Software under the terms of the Free Software Foundation's GNU General Public License in ASCII text file form. It compiles and runs on a wide variety of UNIX platforms and similar systems (including FreeBSD and Linux), Windows and MacOS.

The R environment

R is an integrated suite of software facilities for data manipulation, calculation and graphical display. It includes an efficient data handling and storage facility, a suite of operators for calculations on arrays, especially matrices, a large, coherent, integrated collection of intermediate tools for data analysis, graphical facilities for data analysis and display either on-screen or on hardcopy, and a well-developed, simple and effective programming language which incorporates conditionals, loops, user-defined recursive functions and input and output facilities.

The term "environment" is meant to characterize it as a totally planned and coherent system, instead of an incremental accretion of very specific and inflexible tools, as is usually the case with other data analysis software.

R, like S, is meant around a real computer-oriented language, and it allows users to feature additional functionality by defining new functions. Much of the system is itself written in the R dialect of S, which makes it easy for users to follow the algorithmic choices made. For computationally-intensive tasks, C, C++ and Fortran code are often linked and called at run time. Advanced users can write C code to control R objects directly.

Many users consider R as a statistics system. We like better to consider it of an environment within which statistical techniques are implemented. R can be extended (easily) via packages. There are about eight packages supplied with the R distribution and many more are available through the CRAN family of Internet sites covering a very wide range of modern statistics.

IV. EXPERIMENTAL RESULT

R has its own LaTeX-like documentation format, which is used to supply comprehensive documentation, both on-line in a number of formats and in hardcopy.

summary(paddy)

```

Crop year      Long-grain  Medium-grain
Length:68     Min. :41.60 Length:68
Class :character 1st Qu.:60.63 Class :character
Mode :character  Median :68.87  Mode
:character
                Mean :65.39
                3rd Qu.:73.18
                Max. :79.52
                NA's :14
Short-grain    Total production  Temperature
Min. : 0.4725 Length:68 Length:68
1st Qu.: 1.3138 Class :character Class :character

```

```

Median : 1.9162  Mode :character  Mode
:character
Mean : 4.2678
3rd Qu.: 8.1673
Max. :11.9000
NA's :14
Environment    Disease      Crop year__1
Length:68      Length:68   Length:68
Class :character Class :character Class
:character
Mode :character Mode :character Mode
:character

```

```

United States  Arkansas  California
Min. :3411 Min. :3500 Min. :4325
1st Qu.:4548 1st Qu.:4530 1st Qu.:5658
Median :5555 Median :5300 Median :7380
Mean :5578 Mean :5492 Mean :7018
3rd Qu.:6647 3rd Qu.:6545 3rd Qu.:8225
Max. :7694 Max. :7560 Max. :8890
NA's :9 NA's :9 NA's :9
india banglore kerala
Length:68 Min. :2950 Min. :3300
Class :character 1st Qu.:4275 1st Qu.:4373
Mode :character Median :5400 Median :5100
                Mean :5371 Mean :5242
                3rd Qu.:6650 3rd Qu.:6265
                Max. :7420 Max. :7770
                NA's :9 NA's :9

```

```

up
Min. :2900
1st Qu.:4580
Median :5600
Mean :5631
3rd Qu.:6770
Max. :8370
NA's :9

```

Knn algorithm:

Step1:

```

str(paddy)
Classes 'tbl_df', 'tbl' and 'data.frame':      68
obs. of 16 variables:
 $ Crop year      : chr NA NA "1965""1966" ...
 $ Long-grain    : num NA NA 43 41.6 48.5 46.8
49 49.3 NA 52.6 ...
 $ Medium-grain  : chr "-----"
- - Percent - - - - -" NA
"45.6""46.5" ...
 $ Short-grain   : num NA NA 11.4 11.9 9.2 11.1
10.7 10.3 NA 10.2 ...
 $ Total production: chr "1,000 cwt" NA
"76281""85020" ...
 $ Temperature   : chr NA NA "23.5° to
35°""23.5° to 35°" ...
 $ Environment   : chr NA NA "warm""warm" ...

```

```
$ Disease      : chr NA NA "Bacterial leaf
blight""Bacterial leaf blight" ...
$ Crop year__1 : chr NA NA NA "1960" ...
$ United States : num NA NA NA 3423 3411 ...
$ Arkansas     : num NA NA NA 3525 3500 ...
$ California   : num NA NA NA 4775 4800 ...
$ india       : chr NA "- - - - -"
- - - - - Pounds per acre - - - - -
- - - - - "|__truncated__ NA "2850" ...
$ banglore    : num NA NA NA 2950 3300
3200 3900 3800 3700 4300 ...
$ kerala      : num NA NA NA 3400 3300 4200
4200 4300 4500 4400 ...
$ up          : num NA NA NA 3075 2900 ...
```

V. CONCLUSION

In this thesis classification of healthy and diseased paddy leaves is identified using kNN classifier with DBSCAN features. There are three types of diseases called viral, bacterial and fungal. In this paper, we have considered only identification of healthy and diseased leaf. classify the accuracy level.

REFERENCE:

- [1]. Ansari, M. M., & Sridhar, R. (2012). Iron nutrition and virulence in *Xanthomonas oryzae*. Indian Phytopathology.
- [2]. Aryal L, Bhattarai. G, Subedi A Subedi .M , Subedi .M & Sah. G.K. (2016, April-June) Response of Rice Varieties to Brown Spot Disease of Rice at Paklihawa, Rupandehi, Global Journal of Biology, Agriculture and Health Sciences, ISSN: 2319-5584, G.J.B.A.H.S., Vol.5(2) pp: 50-54.
- [3]. Caglayan, A., Guclu, O., & Can, A. B. (2013, September). A plant recognition approach using shape and color features in leaf images. In International Conference on Image Analysis and Processing (pp. 161-170). Springer, Berlin, Heidelberg.
- [4]. Datnoff, L. E., & Lentini, R. S. (1994). Brown spot in Florida rice.
- [5]. Daud, S. M., Jozani, H. J., & Arab, F. (2013). A review on predicting outbreak of tungro disease in rice fields based on epidemiological and biophysical factors. International Journal of Innovation, Management and Technology, 4(4), 447.
- [6]. Gnanamanickam, S. S. 2009. Major diseases of rice, in Biological Control of Rice Diseases, Springer publications, ch. 2, pp. 32-34.
- [7]. Guyer, D., Tuttle, A., Rouse, S., Volrath, S., Johnson, M., Potter, S., ... & Ward, E. (1998). Activation of latent transgenes in *Arabidopsis* using a hybrid transcription factor. *Genetics*, 149(2), 633-639.
- [8]. Hibino, H. (1996). Biology and epidemiology of rice viruses. Annual review of phytopathology, 34(1), 249-274.
- [9]. Hibino, H., Roechan, M., & Sudarisman, S. (1978). Association of two types of virus particles with penyakitabang (tungro disease) of rice in Indonesia. *Phytopathology*, 68(10), 1412-1416.
- [10]. Jamal-u-ddin Hajano, M. A., Pathan, Q. A., & Mubeen, A. L. (2011). Rice blast-mycoflora, symptomatology and pathogenicity. International Journal for Agro Veterinary and Medical Sciences, 5(1), 53-63.