PADDY DISEASE DETECTION USING IMAGE PROCESSING AND MACHINE LEARNING ALGORITHM

J.RESHMAN MASUTHA [1]
Research Scholar,
Khadir Mohideen College,
Adirampattinam.

Dr .A.SHAIK ABDUL KHADIR,M.Sc.,M.Phil,Ph.D [2]
Head & Associate Professor Of Computer Science,
Khadir Mohideen College,
Adirampattinam.

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. 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 grisae) Leaf blast is difficult to control because it's ability to forms pathogen strains quickly .

Brown Spot (Helminthosporium oryzae) 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. .

Bacterial blight (BLB) may be a disease caused by the bacterium Xanthomonas campestris.

Tungro (TG) is one of the crucial 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, he loss caused by tungro are going to be greater.

2 .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 Magnaporthegrisea, 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 .

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 attach 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 tackles this problem. One way to identify the diseases in plants is by observing the physical changes (diseases spots or lesions) caused by chemical changes in the sick plants . The images of these spots can be processed and used to

recognize the diseases quickly, easily, and inexpensively . 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.

3.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 ...
```

Table 1:Paddy Disease

Name: Disease Type: Nominal
Missing: 0 (0%) Distinct: 3 Unique: 0 (0%)

No.	Label	Count
1	Bacterial leaf blight	13
2	blast kernal smut	24
3	brownspot	17

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances	47	87.037 %
Incorrectly Classified Instances	7	12.963 %
Kappa statistic	0.8038	
Mean absolute error	0.0901	
Root mean squared error	0.2432	
Relative absolute error	20.8238 %	
Root relative squared error	52.2613 %	
Total Number of Instances	54	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
	1	0.146	0.684	1	0.813	0.996	Bacterial leaf blight
	0.833	0.033	0.952	0.833	0.889	0.981	blast kernal smut
	0.824	0	1	0.824	0.903	0.975	brownspot
Weighted Avg.	0.87	0.05	0.903	0.87	0.875	0.982	

=== Confusion Matrix ===

a b c <-- classified as

13 0 0 | a = Bacterial leaf blight

4 20 0 | b = blast kernal smut

2 1 14 | c = brownspot

Cluster:

=== Run information ===

Scheme:weka.clusterers.SimpleKMeans -N 2 -A "weka.core.EuclideanDistance -R first-last" -I 500 -S 10

Relation: paddy

Instances: 54

Attributes: 15

Crop year

Long-grain

Medium-grain

Short-grain

Total production

Temperature

Environment

Disease

United States

Arkansas

California

india

banglore

kerala

up

Test mode:evaluate on training data

=== Model and evaluation on training set ===

kMeans

======

Number of iterations: 4

Within cluster sum of squared errors: 460.6871106650641

Missing values globally replaced with mean/mode

Cluster centroids:

		Cluster#	
Attribute	Full Data	0	1
	(54)	(34)	(20)
Crop year	1991.5	1994.5	1986.4
Long-grain	65.3926	67.3765	62.02
Medium-grain	30.3444	29.2706	32.17
Short-grain	4.2685	3.3588	5.815
Total production	76,281	89,379	76,281
Temperature	40 øC (104 øF).	40 øC (104 øF).	23.5ø to 35ø
Environment	humidity	dry	warm
Disease	blast kernal smut	blast kernal smut	Bacterial leaf blight
United States	3,423	3,411	3,423

Arkansas	3,525	4,300	3,525
California	4,775	8,500	4,775
india	2,850	3,550	2,850
banglore	2,950	4,300	2,950
kerala	3,400	4,600	3,400
up	3,075	6,000	3,075

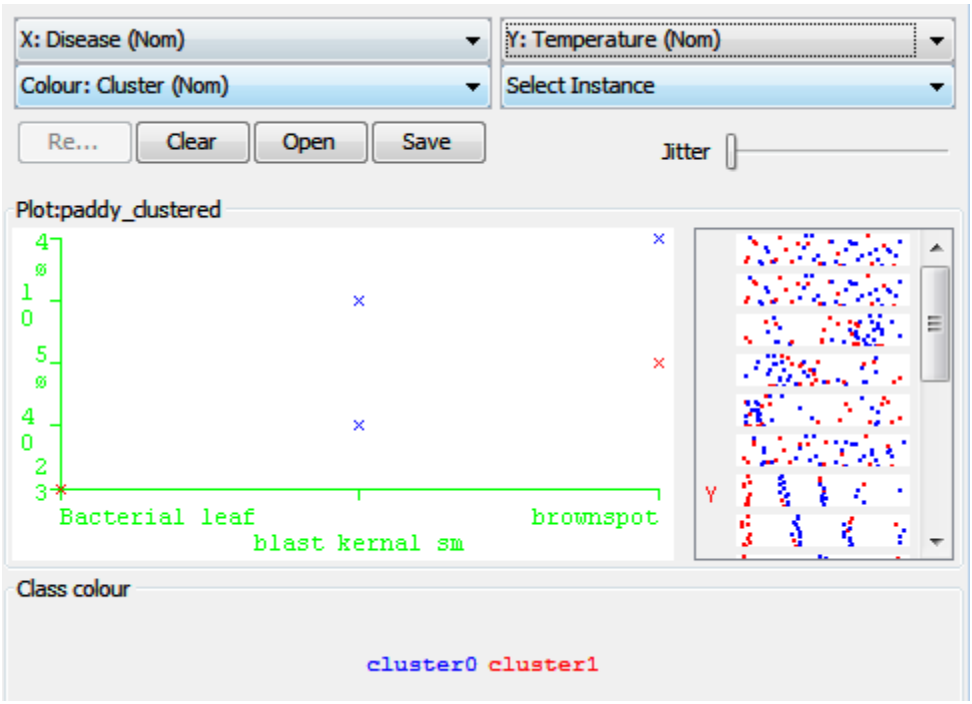
Time taken to build model (full training data) : 0.02 seconds

=== Model and evaluation on training set ===

Clustered Instances

0 34 (63%)

1 20 (37%)



Weka cluster using Paddy

CONCLUSION

In this thesisclassification 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 Xanthomonasoryzaepv. oryzae. Indian Phytopathology.

[2] Aryal L, Bhattarai. G, SubediA 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 penyakithabang (tungro disease) of rice in Indonesia. Phytopathology, 68(10), 1412-1416.

[10] Jamal-u-ddinHajano, 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.