

**PADDY DISEASE DETECTION USING IMAGE PROCESSING**

**BY**

**ABID AL AMIN**

**ID: 151-15-5340**

**RAFEED AN NOOR**

**ID:151-15-5364**

**AND**

**MD. ROKIBUL HASAN**

**ID: 151-15-377**

This Report Presented in Partial Fulfillment of the Requirements for the  
Degree of Bachelor of Science in Computer Science and Engineering

Supervised By

**Md. Azizul Hakim**

**Lecturer**

Department of CSE  
DaffodilInternational University



**DAFFODIL INTERNATIONAL UNIVERSITY**

**DHAKA, BANGLADESH**

**APRIL2019**

## APPROVAL

This Project/internship titled “**Paddy Disease Detection using Image Processing**”, submitted by Abid Al Amin, ID No: 151-15-5340, Md. Rokibul Hasan, ID No: 151-15-377 , Rafeed An Noor, ID No: 151-15-5364 to the Department of Computer Science and Engineering, Daffodil International University has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of B.Sc. in Computer Science and Engineering and approved as to its style and contents. The presentation has been held on 02-05-2019.

## BOARD OF EXAMINERS



**Dr. Syed Akhter Hossain**  
**Professor and Head**

Department of Computer Science and Engineering  
Faculty of Science & Information Technology  
Daffodil International University

**Chairman**



**Dr. Md. Ismail Jabiullah**  
**Professor**

Department of Computer Science and Engineering  
Faculty of Science & Information Technology  
Daffodil International University

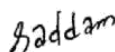
**Internal Examiner**



**Dr. Sheak Rashed Haider Noori**  
**Associate Professor & Associate Head**

Department of Computer Science and Engineering  
Faculty of Science & Information Technology  
Daffodil International University

**Internal Examiner**



**Dr. Md. Saddam Hossain Mukta**  
**Assistant Professor**

Department of Computer Science and Engineering  
United International University

**External Examiner**

## ACKNOWLEDGEMENT

First we express our heartiest thanks and gratefulness to almighty God for His divine blessing makes us possible to complete the final year project/internship successfully.

We really grateful and wish our profound our indebtedness to **Md. Azizul Hakim Lecturer**, Department of CSE Daffodil International University, Dhaka. Deep Knowledge & keen interest of our supervisor in the field of “Image Processing” to carry out this project. His endless patience, scholarly guidance, continual encouragement, constant and energetic supervision, constructive criticism, valuable advice, reading many inferior draft and correcting them at all stage have made possible to complete this project.

We would like to express our heartiest gratitude to **Prof. Syed Akhter Hossain Professor and Head, Department of CSE**, for his kind help to finish our project and also to other faculty member and the staff of CSE department of Daffodil International University.

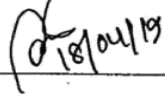
We would like to thank our entire course mate in Daffodil International University, who took part in this discuss while completing the course work.

Finally, we must acknowledge with due respect the constant support and patients our parents.

## DECLARATION

We hereby declare that, this project has been done by us under the supervision of **Md. Azizul Hakim, Lecturer, Department of CSE** Daffodil International University. We also declare that neither this project nor any part of this project has been submitted elsewhere for award of any degree or diploma.

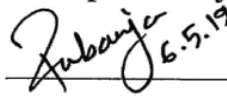
Supervised by:



---

**Md. Azizul Hakim**  
**Lecturer**  
Department of CSE  
Daffodil International University

Co-Supervised by:



---

**Rubaiya Hafiz**  
**Lecturer**  
Department of CSE  
Daffodil International University

Submitted by:

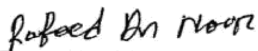


---

**Abid Al Amin**  
ID: 151-15-5340  
Department of CSE  
Daffodil International University

MD. ROKIBUL HASAN

**Md. Rokibul Hasan**  
ID: 151-15-377  
Department of CSE  
Daffodil International University



---

**Rafeed An Noor**  
ID: 151-15-5364  
Department of CSE  
Daffodil International University

## **ABSTRACT**

Being an agricultural country, most of the people of Bangladesh are dependent on agriculture directly or indirectly. It is the fourth largest rice producing country in the world. Main hindrance in rice production is paddy diseases. So in this research the main objective is to develop a prototype system for detecting the paddy diseases, which are Paddy Blast, Brown Spot and Narrow Brown Spot diseases. This concentrate on the image processing techniques used to find pattern in the image and artificial neural network technique to classify the diseases. The methodology involves image collection, image processing, feature extraction and classification. Features are extracted from the images using Haralick's texture feature from color co-occurrence matrix. Then an artificial neural network is trained by these features and a trained model is found. In testing phase, all paddy samples are passed through the leaf color analysis to detect the normal paddy leaf image. If the sample passes leaf color analysis, then it is automatically classified as Normal Paddy leaf image. Otherwise, all the segmented paddy disease samples are converted into the features data and is passed through the artificial neural network. Consequently, by employing the artificial neural network technique, the paddy diseases are recognized. The accuracy to detect diseases of this model is good enough to use in practical life.

## TABLE OF CONTENTS

<b>CONTENTS</b>	<b>PAGE</b>
Board of examiners	i
Declaration	ii
Acknowledgements	iii
Abstract	iv
<b>CHAPTER</b>	
<b>CHAPTER 1: INTRODUCTION</b>	<b>1-3</b>
1.1 Introduction	1
1.2 Background	1
1.3 Problem Statement	2
1.4 Expected Output	3
1.5 Scope of Study	3
1.6 Thesis Organization	3
<b>CHAPTER 2: LITERATURE REVIEW</b>	<b>4-5</b>
2.1 Introduction	4
2.2 Paddy Overviews	4-5
2.3 Case Study on Existing System	5
2.4 Conclusion	5
<b>CHAPTER 3: PROPOSED METHODOLOGY</b>	<b>6-18</b>
3.1 Introduction	6
3.2 Workflow	6

3.3 Image Collection	6-7
3.4 Image Processing	7-8
3.5 Feature Extraction	8-16
3.6 Classification	16-18
<b>CHAPTER 4: EXPERIMENTAL ANALYSIS</b>	<b>19-24</b>
4.1 Introduction	19
4.2 Experimental Setup	19
4.3 Performance Measures – Definition	20
4.4 Result Analysis	21-24
4.5 Conclusion	24
<b>CHAPTER 5: CONCLUSION</b>	<b>25-26</b>
5.1 Summary	25
5.2 Limitation	25
5.3 Future Scope of Work	26
<b>APPENDIX</b>	<b>27</b>
Appendix A: Research Reflection	
<b>REFERENCES</b>	<b>28</b>

## LIST OF FIGURES

<b>FIGURES</b>	<b>PAGE NO</b>
Figure 1.1: Rice Production and Consumption Statistics in Bangladesh	2
Figure 3.1: Work Flow of proposed methodology	6
Figure 3.2: Sample of collected images	7
Figure 3.3: L*a*b* Color Space in Dimensional Graph	7
Figure 3.4: Processing image from RGB to Lab	8
Figure 3.5: Brown Spot affected paddy image	11
Figure 3.6: Proposed Artificial Neural Network	18
Figure 4.1: Correctly classified 4 Paddy Blast images	22
Figure 4.2: Incorrectly classified 4 Paddy Blast images	22
Figure 4.3: Correctly classified 4 Brown Spot images	23
Figure 4.4: Incorrectly classified 4 Brown Spot images	23
Figure 4.5: Correctly classified 4 Narrow Brown Spot images	23
Figure 4.6: Incorrectly classified 4 Narrow Brown Spot images	24
Figure: Plagiarism Report	



## LIST OF TABLES

<b>TABLES</b>	<b>PAGE NO</b>
Table 3.1:A 20x20 matrix of CCM values	14
Table 4.1:Confusion matrix of proposed system	23
Table 4.2:Accuracy, Precision and Recall with k-fold(5-fold) method	24

# CHAPTER 1

## Introduction

### 1.1 Introduction

This chapter discuss about an overview of the study conducted. The title is “Paddy Disease Detection using Image Processing”. This papercontains background, problem statements, objectives and the scope of study. The background briefly describes the identification of thesis and related issues. Problem statements describes the problems that arise and make the selected projects to be undertaken. The objectives are the goals list for the research to be achieved. Scope of study discuss about the limitations of research. Lastly, thesis organization gives a summary of the sequence for each chapter in the paper.

### 1.2 Background

Rice known as *Oryza Sativa* (specific name), is one of the most utilized food plants and widely grown originated in ASIA. Rice is an important crop worldwide and over half of the world population relies on it for food. Many people in the world including Bangladesh eat rice as staple food. However, there are many factors that make paddy rice production become slow and less productive. One of the main factors is paddy disease.

An abnormal condition that injures the plant or leads it to function improperly is called as a disease. Diseases are readily recognized by their symptoms. There are a lot of paddy disease types which are paddy blast, narrow brown spot, brown spot disease and many more. Image processing and computer vision technology are very beneficial to the agricultural industry. They are more potential and more important to many areas in agricultural technology.

As an agricultural country, Bangladesh gets its one-sixth of national income from rice. About 10.5 million hectares’ lands produce 25.0 million tons’ rice ever year [2014]. Now

govt.'s target is to produce another 30 million over the next 20 years. Main Obstacle for gaining the target is those paddy diseases. If diseases can detect easily with image processing, taking action will be faster. In the fig 1.1 the red line represents annual consumption and the blue line represent annual production from year 1961 to 2006.

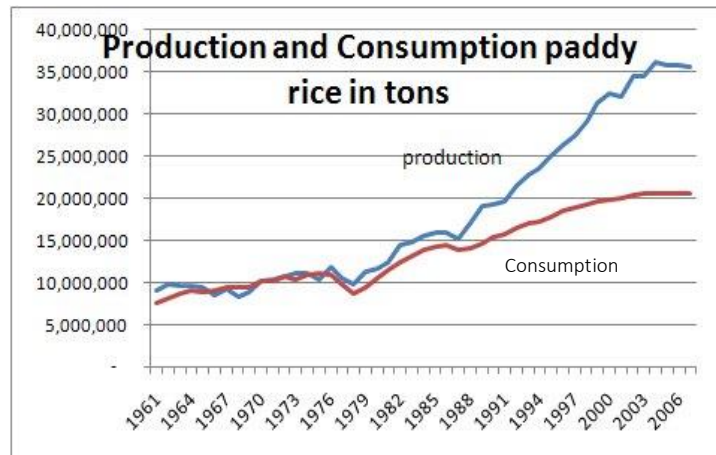


Figure 1.1: Rice Production and Consumption Statistics in Bangladesh

Paddy Disease Detection System is one of the very beneficial systems. It can help the paddy farmer detect the disease faster. This study aims to develop a prototype system to automatically detect and classify the paddy diseases by using image processing technique as an alternative or supplemental to the traditional manual method.

### 1.3 Problem Statement

Paddy will be harvest twice in a year. Most of paddy farmer faces many problems to harvest their paddy because they used to attack by snail, worm and fungi. Furthermore, when the paddy had been infected or attacked, the others areas had been exposed to be infected. Thus, it will decrease paddy farmer's income and lead to significance losses to farmer. Currently, the paddy farmer determines the type of disease manually. The errors might occur in order to determine the type of diseases. Paddy farmer also have to spend a lot of time to detect the type of disease. It also takes a time as the paddy farmers manually check the disease since the paddy field is in wide area.

#### **1.4 Expected Output**

The development of the application has met its specification successfully. For the given leaves with three categories of disease, the application can recognize all them.

#### **1.5 Scope of Study**

The users of the system are general farmer. The prototype is designed in python and 220 total samples of normal, brown spot disease, narrow brown spot disease and blast disease is used in this.

#### **1.6 Thesis Organization**

This thesis consists of 5 chapters ranging from Chapter 1 until Chapter 5. Chapter 1 gives an overview of the topic. It additionally comprises of Problem Statements, Expected Output and the Scope of Study. In the meantime, Chapter 2 discusses the past research works that was led by others explores. All the significant specialized paper, diaries and books taken from those inquires about will be talked about in detail. Chapter 3 centers about the procedure for framework improvement and process stream in detail of this examination. It uncovers the method and the calculations that will be utilized as a part of playing out this investigation. Chapter 4 that comprise of expected outcome or yield, limitation of task and further research. In conclusion, Chapter 5 finishes up the general examination, research and limitation.

## **CHAPTER 2**

### **Literature Review**

#### **2.1 Introduction**

This chapter quickly surveys, clarifies and talks about on existing writing audit related with our research topic which is "Study on Paddy Disease Detection using Color Co-occurrence Features". This part includes three areas. The primary segment portrays the outlines of paddy. The subsections are the definition, kind of paddy infection, paddy manifestation and paddy administration. The second area is the survey of some current framework that utilized same strategies and techniques. The third area talks about the survey on strategy and technique utilized by the framework. The subsections are picture securing, picture division and manufactured neural system.

#### **2.2 Paddy Overviews**

In this area, firstly introduces a meaning of paddy. From that point onward, this subsection quickly examines onsymptoms of paddy diseases.

##### **2.2.1 Definition of Paddy**

Paddy otherwise called rice is the dull seeds of a yearly south-east Asian grain grass (*Oryza sativa*) that are cooked and utilized for sustenance. This grain grass that is broadly developed in warm atmospheres for its seeds and results. Rice is a standout amongst the most used sustenance plants and generally developed began in ASIA. Rice is a critical product worldwide and over portion of the total populace depends on it for sustenance. Numerous individuals on the planet including Bangladesh eat rice as staple nourishment.

## **2.2.2 Paddy Diseases Symptoms**

### **2.2.2.1 Paddy Blast Symptoms**

### **2.2.2.2 Brown Spot Symptoms**

### **2.2.2.3 Narrow Brown Spot Symptoms**

- 

## **2.3 Case Study on Existing System**

## **2.4 Conclusion**

In this chapter, this thesis presented the common method for detecting paddy disease. By previewing various method, we learned advantage and disadvantage of various method and their characteristic for simple and clean paddy disease detection.

## CHAPTER 3

### Proposed Methodology

#### 3.1 Introduction

The purpose of this chapter is to discuss the approach and system for the thesis. Method, technique or approach that has been used while designing and implementing the thesis are included in this chapter. This chapter also explains about the justification of method or approach used and hardware and software necessity.

#### 3.2 Work Flow

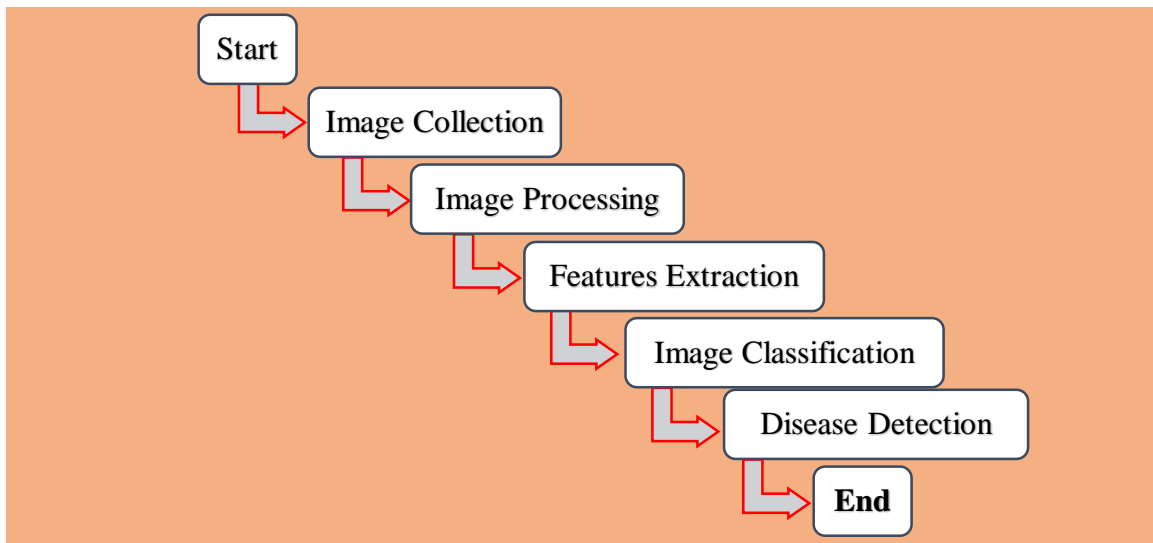
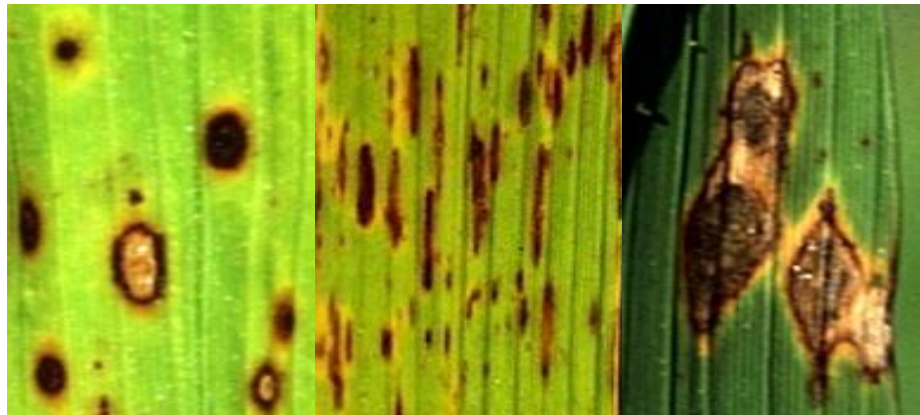


Figure 3.1: Work Flow of proposed methodology

#### 3.3 Image Collection

The RGB images of paddy leaf are collected from Internet. Those image cropped into a smaller image with dimension of 64 x 64 pixels as training data. We have collected about 180 data samples with the four rotation from each images. It consists of three types of paddy diseases (Paddy Blast, Brown Spot, Narrow Brown Spot) as shown in Fig. 3.2. Images are stored in jpg format.



(a) (b) (c)

Figure 3.2: Sample of collected images (a) Brown Spot Disease; (b) Narrow Brown Spot Disease; (c) Blast Disease

### 3.4 Image Processing

The main objective of this process is to obtain an image with an approximation of human color perception. The RGB image (Fig. 3.4(a)) is converted into Lab as abbreviation for CIE  $L^*a^*b^*$  1976 color space (also CIELAB), as shown in Fig. 3.4(b).

#### Lab Color Space

The LAB color model is a three axis color system and LAB colors are absolute meaning that the color is exact. It's what's known as device independent meaning that the LAB color space is the only way to communicate different colors across different devices. An object's color is measured in LAB color with a spectrophotometer.

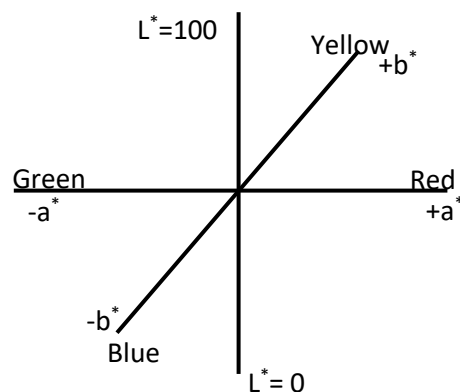
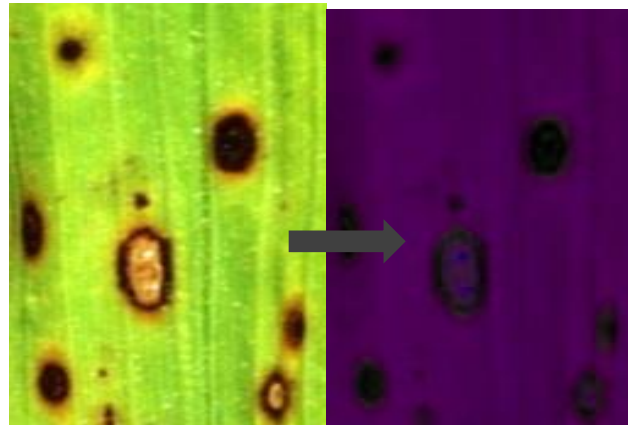


Fig 3.3:  $L^*a^*b^*$  Color Space in Dimensional Graph





(a) (b)

Figure 3.4: Processing image from RGB to Lab (a) RGB image; (b) Lab image

### 3.5 Feature Extraction

Feature Extraction a sort of dimensionality lessening that productively speaks to intriguing parts of a picture as a smaller component vector. Features are extracted from the color co-occurrence matrix which is calculated previously.

### 3.5.1 Color Co-occurrence Matrix

A co-occurrence matrix or co-occurrence distribution is a matrix that is defined over an image to be the distribution of co-occurring pixel values (grayscale values, or colors) at a given offset:

- The offset,  $(\Delta x, \Delta y)$ , is a position operator that can be applied to any pixel in the image (ignoring edge effects): for instance,  $(1, 2)$  could indicate "one down, two right".
- An image with  $p$  different pixel values will produce a  $p \times p$  co-occurrence matrix, for the given offset.
- The  $(i, j)^{th}$  value of the co-occurrence matrix gives the number of times in the image that the  $i^{th}$  and  $j^{th}$  pixel values occur in the relation given by the offset.

$$C_{\Delta x, \Delta y}(i, j) = \sum_{x=1}^n \sum_{y=1}^m \begin{cases} 1, & \text{if } I(x, y) = i \text{ and } I(x + \Delta x, y + \Delta y) = j \\ 0, & \text{otherwise} \end{cases}$$

Where:  $i$  and  $j$  are the pixel values;  $x$  and  $y$  are the spatial positions in the image  $I$ ; the offsets  $(\Delta x, \Delta y)$  define the spatial relation for which this matrix is calculated; and  $I(\Delta x, \Delta y)$  indicates the pixel value at pixel  $(x, y)$ .

The offset value  $(\Delta x, \Delta y)$  is calculated by the spatial direction. If the direction is

$$0^\circ, \text{ then } \Delta x = 0, \Delta y = 1$$

$$45^\circ, \text{ then } \Delta x = 1, \Delta y = 1$$

$$90^\circ, \text{ then } \Delta x = 1, \Delta y = 0$$

$$135^\circ, \text{ then } \Delta x = -1, \Delta y = 1$$

Let's take an example of an  $5 \times 5$  image and the GLCM matrix is calculated by following procedure:

$$\begin{bmatrix} 22 & 220 & 180 & 111 & 75 \\ 25 & 240 & 103 & 180 & 118 \\ 65 & 110 & 210 & 230 & 191 \\ 123 & 40 & 150 & 5 & 199 \\ 180 & 95 & 70 & 15 & 255 \end{bmatrix}$$

The matrix size is  $5 \times 5$ . So, the segment size will be  $255/5 = 51$  and all values of the above matrix will be replaced by below:

$$\begin{aligned} 0 - 51 &\rightarrow 0 \\ 52 - 103 &\rightarrow 1 \\ 104 - 154 &\rightarrow 2 \\ 155 - 205 &\rightarrow 3 \\ 206 - 255 &\rightarrow 4 \end{aligned}$$

By applying above mapping the resultant matrix is given below:

$$\begin{bmatrix} 04321 \\ 04132 \\ 12443 \\ 20203 \\ 31104 \end{bmatrix}$$

When the direction is  $0^\circ$  and the offset value is 1, then the matrix which will be found is given below:

$$\begin{bmatrix} 00113 \\ 11110 \\ 21001 \\ 01200 \\ 01021 \end{bmatrix}$$

The transpose matrix of the above matrix is given below:

$$\begin{bmatrix} 01200 \\ 01111 \\ 11020 \\ 11002 \\ 30101 \end{bmatrix}$$

After Adding the above two matrices, we get the following matrix:

$$\begin{bmatrix} 01313 \\ 12221 \\ 32021 \\ 12202 \\ 31122 \end{bmatrix}$$

Now, the determinant of this matrix = 40. By normalizing the matrix, we get resultant matrix:

$$\begin{bmatrix} 0 & 0.025 & 0.075 & 0.025 & 0.075 \\ 0.025 & 0.050 & 0.050 & 0.050 & 0.025 \\ 0.075 & 0.050 & 0 & 0.050 & 0.025 \\ 0.025 & 0.050 & 0.050 & 0 & 0.050 \\ 0.075 & 0.025 & 0.025 & 0.050 & 0.050 \end{bmatrix}$$

This is the resultant GLCM matrix when the offset value is 1 and rotation is 0°.

If we apply CCM in the Fig 3.5 image of size 64 x 64, a 20 x 20CCM matrix will be found.



Figure 3.5: Brown Spot affected paddy image

Table 3.1: A 20x20 matrix of CCM values

105	52	25	8	4	4	4	3	4	3	3	3	2	2	2	1	0	1	0	0
42	50	38	20	13	9	8	4	2	2	3	4	2	1	3	1	1	1	0	0
26	31	39	31	17	14	7	7	5	2	2	2	2	3	1	2	2	2	0	0
9	16	27	42	26	15	12	12	4	6	4	4	4	4	2	4	4	2	0	0
8	12	16	22	30	23	13	11	10	4	5	4	2	5	3	3	1	2	0	0
5	11	13	15	30	35	25	17	14	6	7	7	4	4	2	3	4	1	1	0
7	7	8	11	14	27	47	37	19	11	8	6	5	5	5	3	2	3	0	0
5	3	7	10	9	22	39	55	28	18	15	9	6	5	10	4	1	2	5	1
3	3	7	6	15	21	37	42	23	19	13	9	6	4	5	2	4	1	2	3
2	2	3	4	5	7	10	17	25	25	23	17	10	6	4	5	4	3	3	0
1	6	2	5	5	8	7	14	21	22	29	28	18	11	7	4	4	4	2	1
5	4	3	3	3	10	9	9	14	15	26	31	26	18	15	8	9	7	3	2
3	3	4	3	3	3	5	5	10	7	19	26	23	21	22	12	8	5	4	1
2	2	2	4	4	6	7	6	6	6	12	19	21	26	30	18	10	8	4	2
2	2	1	3	2	3	4	8	7	6	7	17	20	30	29	23	18	14	8	6
1	1	1	3	2	3	4	5	5	5	5	9	11	18	28	34	24	17	12	3
0	2	1	2	0	3	3	4	3	5	4	11	10	11	18	24	28	27	13	5
1	1	1	1	1	1	2	1	2	3	4	4	5	8	9	13	20	30	49	35
0	0	0	0	0	0	1	3	2	3	3	4	4	6	9	10	14	34	43	14
0	0	0	0	0	0	0	0	1	2	1	1	2	2	3	3	6	12	14	4

### 3.5.2 Textural Features

Notations:

$p(i, j)$   $(i, j)$ th entry in the normalized gray level co-occurrence matrix

$p_x(i)$   $i$ th entry in the marginal probability matrix obtained by summing the rows of  $p(i, j)$

$N_g$  number of distinct gray levels in the quantized image.

1) Angular Second Moment:

$$f_1 = \sum_i \sum_j (p(i, j))^2$$

Angular Second Moment measure the smoothness of the image. There are two cases,

If all pixels have same gray level  $I=k$ ,

$$p(k, k) = 1 \text{ if } (i = j) \text{ and } p(i, j) = 0 \text{ if otherwise.}$$

$$ASM = 1$$

If all pixels have different gray level,

$$p(i, j) = 1/R \text{ \& } ASM = 1/R$$

ASM value of Fig 3.5 image is: 0.00528

2) Contrast:

$$f_2 = \sum_{n=0}^{N_g-1} n^2 \left\{ \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i, j) \right\}$$

Contrast measures the image contrast (locally gray level variations). The term  $n^2$  is used to take of the largest contrast value.

Contrast value of Fig 3.5 image is: 26.67

3) Correlation:

$$f_3 = \frac{\sum_i \sum_j (ij) p(i, j) - \mu_x \mu_y}{\sigma_x \sigma_y}$$

Correlation measures how the pixels are correlated with each other. Where  $\mu_x, \mu_y$  are the standard deviation and  $\sigma_x, \sigma_y$  are means of  $p_x, p_y$

Correlation value of Fig 3.5 image is: 0.557

4) Sum of squares: Variance

$$f_4 = \sum_i \sum_j (i - \mu)^2 p(i, j)$$

Sum of squares value of Fig 3.5 image is: 128.231

5) Inverse Difference Moment (Homogeneity)

$$f_5 = \sum_i \sum_j \frac{1}{1 + (i - j)^2} p(i, j)$$

Inverse Difference Moment takes care of low contrast images. It takes care of low contrast images because of the inverse  $(i - j)^2$ .

Homogeneity value of Fig 3.5 image is: 0.30597

6) Sum Average

$$f_6 = \sum_{i=2}^{2N_g} i p_{x+y}(i)$$

Sum Average value of Fig 3.5 image is: 19.807

7) Sum Variance

$$f_6 = \sum_{i=2}^{2N_g} (i - f_8)^2 p_{x+y}(i)$$

Sum Variance value of Fig 3.5 image is: 426.844

8) Sum Entropy

$$f_8 = - \sum_{i=2}^{2N_g} p_{x+y}(i) \log\{p_{x+y}(i)\}$$

Sum Entropy value of Fig 3.5 image is: 1.554

9) Entropy

$$f_9 = - \sum_i \sum_j p(i, j) \log\{p(i, j)\}$$

Entropy takes low values for smooth images. It measures the randomness.

Entropy value of Fig 3.5 image is: 2.4052

10) Difference Variance

$$f_{10} = \text{variance of } p_{x-y}$$

Difference Variancevalue of Fig 3.5 image is: 0.00436

11) Difference Entropy

$$f_{11} = - \sum_{i=0}^{N_g-1} p_{x-y}(i) \log\{p_{x-y}(i)\}$$

Difference Entropyvalue of Fig 3.5 image is: 0.99

12) &

13) Information Measure of Correction

$$f_{12} = \frac{HXY - HXY1}{\max\{HX, HY\}}$$

$$f_{13} = (1 - \exp[-2.0(HXY2 - HXY)])^{1/2}$$

$$HXY = - \sum_i \sum_j p(i) \log\{p(i)\}$$

Since some of the probabilities becomes zero and  $\log(0)$  is very high so arbitrary small positive constant is added to avoid the infinite number.

Where,  $HX$  and  $HY$  are entropies of  $p_x$  and  $p_y$  and

$$HXY1 = - \sum_i \sum_j p(i, j) \log\{p_x(i)p_y(j)\}$$

$$HXY2 = - \sum_i \sum_j p_x(i)p_y(j) \log\{p_x(i)p_y(j)\}$$

Information Measure of Correlation 1 value of Fig 3.5 image is: -0.1358

Information Measure of Correlation 2 value of Fig 3.5 image is: 0.4915

14) Maximal Correction Coefficient(Energy)

$$f_{14} = (\text{second largest eigenvalue of } Q)^{1/2}$$

Where,

$$Q(i, j) = \sum_k \frac{p(i, k)p(j, k)}{p_x(i)p_y(k)}$$

Energy value of Fig 3.5 image is: 0.0711



### 3.5.3 Feature Selection

Selecting correct features is an important issue for the system. Not all the features are suitable for classifying the different classes. In our study around all 14 features we found only 5 of them are appropriate. For selecting those features a method called Subset Choosing method is used.

The methodology is:

*ans = 0.0*

*subset = 1*

*for i in range (1, (1<<15)):*

*ret = fun(i)*

*if(ret>ans):*

*ans = ret*

*subset = i*

There are  $2^{15}$  possible subsets among them the subset which is the possible best subset. So, the selected features are:

- 1) Homogeneity
- 2) Angular Second Moment(ASM)
- 3) Energy
- 4) Information Measure of Correlation 1
- 5) Information Measure of Correlation 2

### 3.6 Classification

After extracting features from the images, now a classifier is needed to classify the images. Here an artificial neural network with three hidden layer is used as a classifier. In

the classifier two steps are followed. If an image does not pass leaf color analysis, classifier algorithm will be used to detect the diseases.

### 3.6.1 Leaf Color Analysis

First the whole image is scanned through and calculate the minimum and maximum value for each channel. The RGB calculation will be passed.

$$93 \leq R_{min} \leq 211 \& 93 \leq R_{max} \leq 211$$

$$142 \leq G_{min} \leq 222 \& 142 \leq G_{max} \leq 222$$

$$64 \leq B_{min} \leq 155 \& 64 \leq B_{max} \leq 155$$

If an image passes all the above conditions, then the image is normal leaf image. Otherwise, it is an affected image.

To calculate those values first we took some normal paddy leaf images and calculate the histogram values. From those histogram values the minimum and maximum values were calculated.

### 3.6.2 Artificial Neural Network

All the selected features are in the input of an Artificial Neural Network and the output is used in classification. The equation which is minimized is

$$WX + B$$

Where,

$$W = \text{weights}$$

$$B = \text{biases}$$

$$X = \text{Input Features}$$

The above equation is minimized such that the cost function produces as small error as possible.

In an artificial neural network, there are some nodes and the nodes are connected by some edges. The edges represent some arbitrary values and the nodes represent some activation

functions. In our proposed methodology, Tensor Flow is used as the deep learning framework.

TensorFlow is a typical computational graph that can be executed much more efficiently than if the same calculations were to be performed directly in python. TensorFlow can be more efficient than NumPy because TensorFlow knows the entire computation graph that must be executed, while NumPy knows only the computation of a single mathematical operation at a time. TensorFlow can also automatically calculate the gradients that are needed to optimize the variables of the graph so as to make the model perform better. This is because the graph is a combination of simple mathematical equations. So, the gradient of the entire graph can be calculated through iterative chain rule method.

Fig 3.6 shows our used system, where the network architecture 15 – 50 – 50 – 50 – 4 nodes in 1 input, 3 hidden and 1 output layer respectively. It is a fully connected artificial neural network. Each hidden layer has 50 nodes which gives better result in our research. The hidden nodes are chosen by comparative study.

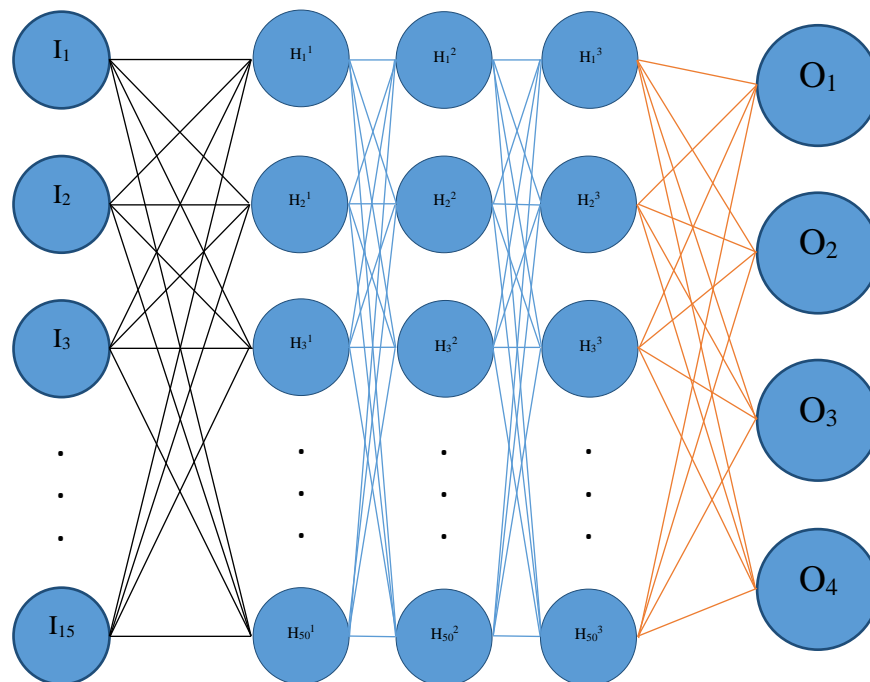


Figure 3.6: Proposed Artificial Neural Network

## **CHAPTER 4**

### **Experimental Analysis**

#### **4.1 Introduction**

There are 220 samples of paddy image used as sample data in the testing phase of this development. The paddy images samples had gone through the phases as discuss in the chapter 3. This chapter will briefly describe about the output result of each phases.

#### **4.2 Experimental Setup**

The experiments and related analysis are done in this chapter. The experiments and analysis processes are done on a computer with Core-I5 processor having 4 cores with each core having 2.5GHz Speed. Also the system had 4GB of RAM, and 1GB of internal intel HD video memory. For software, PyCharm Community Edition 2017.1 is used and program is done with python language with OpenCV and deep learning framework, TensorFlow. OpenCV (Open Source Computer Vision) is a library of programming functions mainly aimed at real-time computer vision. The reason for using OpenCV because it gives easy functionality to do different processes without going into implementations. Moreover, it gives the benefit to use GPU by which processes can be made faster than using CPU only for computation works.

### 4.3 Performance Measures – Definition

### 4.4 Result Analysis

Let's indicate Paddy Blast, Brown Spot, Narrow Brown Spot and Other, 4 classes as PB, BS, NBS and O respectively. And predicted class and actual class are on column and row. So, the table of confusion matrix is as below:

Table 4.1: Confusion matrix of proposed system

		Predicted Class			
		PB	BS	NBS	O
Actual Class	PB	17	1	0	1
	BS	0	13	0	1
	NBS	0	1	7	1
	O	1	0	1	3

#### 4.4.1 Random Test(k-fold)

Random test is used to determine the robustness of a system. Here all the data are taken and from the data randomly some data are selected for testing and some data are selected for training. Here k-fold means, the whole data are divided into k-fold, then each fold is used as testing data and rest of the data are used as training data. And there are n-number of runs.

5-fold is used within 3 distinct run with our collected data and the resulted values are showed in the table 4.1 with the accuracy, precision and recall individually, as average and total.

Table 4.2: Accuracy, Precision and Recall with k-fold(5-fold) method

Run	Fold	Fold Accuracy(%)	Fold Precision(%)	Fold Recall(%)	Run Accuracy(%)	Run Precision(%)	Run Recall(%)	Total Accuracy(%)	Total Precision(%)	Total Recall(%)
1	1	86.4	90.3	73.8	86.38	87.52	80.02	85.69	83.36	76.17
	2	84.1	85.6	78.3						
	3	81.8	78.9	83.3						
	4	93.2	94.2	88.0						
	5	86.4	88.6	76.7						
2	1	84.1	73.4	75.0	84.38	78.72	73.28			
	2	86.4	66.9	70.3						
	3	90.0	92.4	80.5						
	4	77.3	78.3	63.6						
	5	84.1	82.0	77.0						
3	1	79.5	76.7	68.9	86.32	84.74	75.2			
	2	86.3	89.6	74.4						
	3	88.6	89.8	73.9						
	4	88.6	91.3	81.2						
	5	88.6	76.3	77.6						

#### 4.4.2 Classification Result

##### 4.4.2.1 Paddy Blast

Total 87 leaves which are affected by Paddy Blast disease are used for experiment. In those leaves, 10 images were misclassified and the other 77 images are correctly classified.

Examples:

### Correctly classified



Figure4.1: Correctly classified 4 Paddy Blast images

### Incorrectly classified



Figure 4.2: Incorrectly classified 4 Paddy Blast images

Accuracy of Paddy Blast = 88.51%

#### 4.4.2.2 Brown Spot

Total 56 leaves which are affected by Paddy Blast disease are used for experiment. In those leaves, 14 images were misclassified and the other 42 images are correctly classified.

Examples:

### Correctly classified



Figure 4.3: Correctly classified 4 Brown Spot images

**Incorrectly classified**

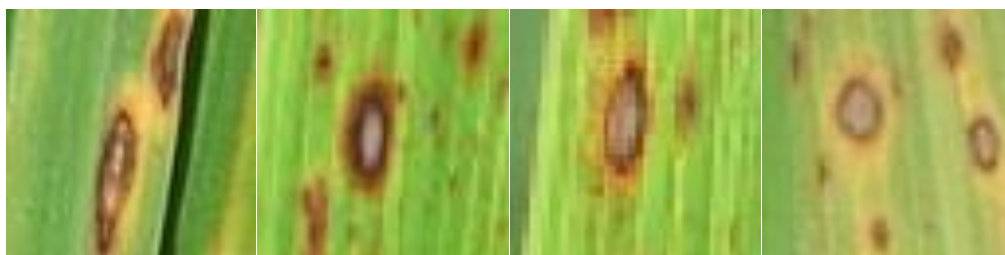


Figure 4.4: Incorrectly classified 4 Brown Spot images

Accuracy of Brown Spot = 75.00%

**4.4.2.3 Narrow Brown Spot**

Total 33 leaves which are affected by Paddy Blast disease are used for experiment. In those leaves, 4 images were misclassified and the other 29 images are correctly classified.

Examples:

**Correctly classified**

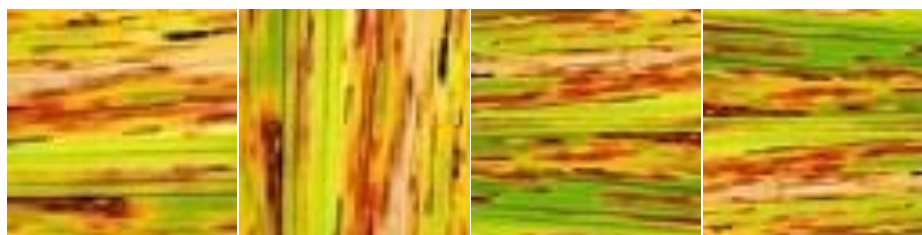


Figure 4.5: Correctly classified 4 Narrow Brown Spot images

**Incorrectly classified**



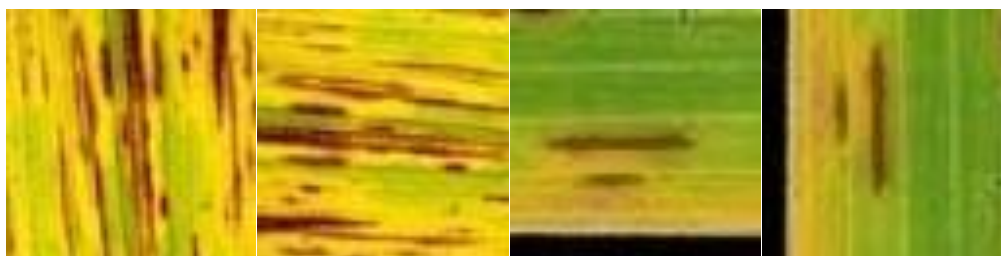


Figure 4.6: Incorrectly classified 4 Narrow Brown Spot images

Accuracy of Narrow Brown Spot = 87.87%

#### **4.5 Conclusion**

In this chapter, a system for diagnosis the paddy disease has been developed using the PyCharm

application. The image processing techniques are applied to find the features and neural network is used to build a trained model. In testing phase, the model is tested with the test images. Before testing each image is passed through leaf color analysis to find out the normal leaf images. The accuracy of testing around 85 percent which is good for practical use.

## **CHAPTER 5**

### **Conclusion**

#### **5.1 Summary**

A system for diagnosis the paddy disease has been developed using the PyCharm application. The image processing techniques is applied to improve and enhance the image to a better quality. Besides, the artificial neural network and leaf color analysis is used to classify the paddy diseases which are paddy blast, brown spot disease, narrow brown spot disease and normal paddy leaf. The methodology involves image collection, image processing, analysis and classification of the paddy disease. All the paddy sample will be passing through the leaf color analysis before it proceeds to the artificial neural network. If the sample is in the range of normal paddy RGB, then it is automatically classifying as Normal paddy leaf and otherwise all the disease affected sample will be passed through image processing to get the features which will be forwarded to the artificial neural network for training and testing. Consequently, by employing the artificial neural network technique, the paddy diseases are recognized about 86 percent accuracy rates. This study has a very great potential to be further improved in the future.

#### **5.2 Limitation**

There are not many impediments in our proposed strategy. One of them is accuracy percentage. The accuracy is at least 86 percent that is not high enough. For better result this number should be risen. There may be also some implementation bug to fix. Best features selection is computationally hard and lengthy process, so, we could not check all possible subset of the features. We only checked some random features based on previous best features. There is a possibility of losing some best features. And also it takes a lot of time, so, our implementation is little slower. Thus it can't be applied on real time for now.

### **5.3 Future Scope of Work**

Even though, we've tried our best to get the desired output and the accuracy value is quite good for this method, yet there is still room for improvement as long as it is not close to 100 percent.

- By using more efficient way to choose best features, this method can be made to work in real time
- Can be implemented in LBP which could rise the accuracy
- Can be implement in mobile application thus farmers can easily detect the disease before it is too late.

## APPENDIX

### Research Reflection

The purpose of this appendix is about research reflection. From Spring 2018 semester we started our journey to make this Application. The main feature of our application is detecting paddy disease. Also using our application people can easily identify paddy disease. In the very first days, we studied a lot of papers about paddy disease and others. We have studied a lot to detect the disease. However, we worked hard and tried our best to learn and research some image processing and machine learning technique so that the application was implemented which achieved the goals in the specification. On the other hand, with few requirement and constraints, the goal is clear and the system design is flexible.

Firstly, we built a model for our project. Then we implemented our project step by step. After many hard works and spending a lot of time finally we were able to reach our goal. The project "Paddy Disease Detection Using Image Processing" will be very helpful for every normal people and farmers. People will be satisfied very easily. So we believe that our "Paddy Disease Detection Using Image Processing" application will be very positive and effective for users.

## REFERENCES

- [1] M. R. Tejonidhi, B. R. Nanjesh, J. G. Math and A. G. D'sa, "Plant disease analysis using histogram matching based on Bhattacharya's distance calculation," *2016 International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT)*, Chennai, 2016, pp. 1546-1549.
- [2] S. Mutalib, M. H. Abdullah, S. Abdul-Rahman and Z. A. Aziz, "A brief study on paddy applications with image processing and proposed architecture," *2016 IEEE Conference on Systems, Process and Control (ICSPC)*, Bandar Hilir, 2016, pp. 124-129.
- [3] R. P. Narmadha and G. Arulvadivu, "Detection and measurement of paddy leaf disease symptoms using image processing," *2017 International Conference on Computer Communication and Informatics (ICCCI)*, Coimbatore, 2017, pp. 1-4.
- [4] N. N. Kurniawati, S. N. H. S. Abdullah, S. Abdullah and S. Abdullah, "Investigation on Image Processing Techniques for Diagnosing Paddy Diseases," *2009 International Conference of Soft Computing and Pattern Recognition*, Malacca, 2009, pp. 272-277.
- [5] R. M. Prakash, G. P. Saraswathy, G. Ramalakshmi, K. H. Mangaleswari and T. Kaviya, "Detection of leaf diseases and classification using digital image processing," *2017 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS)*, Coimbatore, 2017, pp. 1-4.
- [6] P. Revathi and M. Hemalatha, "Classification of cotton leaf spot diseases using image processing edge detection techniques," *2012 International Conference on Emerging Trends in Science, Engineering and Technology (INCOSET)*, Tiruchirappalli, Tamilnadu, India, 2012, pp. 169-173.
- [7] D. Al Bashish, M. Braik and S. Bani-Ahmad, "A framework for detection and classification of plant leaf and stem diseases," *2010 International Conference on Signal and Image Processing*, Chennai, 2010, pp. 113-118.
- [8] A. N. I. Masazhar and M. M. Kamal, "Digital image processing technique for palm oil leaf disease detection using multiclass SVM classifier," *2017 IEEE 4th International Conference on Smart Instrumentation, Measurement and Application (ICSIMA)*, Putrajaya, 2017, pp. 1-6

# PLAGIARISM REPORT

**Checked By:** DIU Project Report Library

Figure: Plagiarism Report

## Document Viewer

# Turnitin Originality Report

Processed on: 03-Apr-2019 10:24 +06

ID: 1104962907

Word Count: 5439

Submitted: 1

## Paddy Disease By Abid Al Amin

Similarity Index  <h1>25%</h1>	<b>Similarity by Source</b>  Internet Sources: N/A Publications: N/A Student Papers: 25%
--------------------------------------	--

[exclude quoted](#)
[exclude bibliography](#)
[exclude small matches](#)
[download](#)  
[refresh](#)
[print](#)
 mode:

13% match (student papers from 18-Dec-2012)	✕
<a href="#">Submitted to Universiti Malaysia Pahang on 2012-12-18</a>	
2% match (student papers from 07-Apr-2018)	✕
<a href="#">Submitted to Daffodil International University on 2018-04-07</a>	
1% match (student papers from 05-Nov-2018)	✕
<a href="#">Submitted to Indian Institute of Technology Guwahati on 2018-11-05</a>	
1% match (student papers from 10-Dec-2017)	✕
<a href="#">Submitted to University of Birmingham on 2017-12-10</a>	
1% match (student papers from 01-Apr-2019)	✕
<a href="#">Submitted to Daffodil International University on 2019-04-01</a>	
1% match (student papers from 01-Jun-2016)	✕
<a href="#">Submitted to Universiti Teknologi Malaysia on 2016-06-01</a>	
1% match (student papers from 07-Apr-2018)	✕
<a href="#">Submitted to Daffodil International University on 2018-04-07</a>	
1% match (student papers from 01-Feb-2017)	✕
<a href="#">Submitted to Bangkok University on 2017-02-01</a>	
1% match (student papers from 05-Jan-2015)	✕
<a href="#">Submitted to Rutgers University, New Brunswick on 2015-01-05</a>	
<1% match (student papers from 05-Feb-2014)	✕
<a href="#">Submitted to Upstate Medical University, SUNY on 2014-02-05</a>	
<1% match (student papers from 01-Dec-2009)	✕
<a href="#">Submitted to Universiti Kebangsaan Malaysia on 2009-12-01</a>	

<1% match (student papers from 31-Dec-2017)  
[Submitted to University of East London on 2017-12-31](#)



<1% match (student papers from 19-Apr-2012)  
[Submitted to VIT University on 2012-04-19](#)



<1% match (student papers from 02-Apr-2019)  
[Submitted to Daffodil International University on 2019-04-02](#)



<1% match (student papers from 08-May-2016)  
[Submitted to Sim University on 2016-05-08](#)



<1% match (student papers from 03-Dec-2013)  
[Submitted to Universiti Malaysia Pahang on 2013-12-03](#)



<1% match (student papers from 03-Apr-2018)  
[Submitted to Daffodil International University on 2018-04-03](#)



<1% match (student papers from 30-Apr-2006)  
[Submitted to UT, Dallas on 2006-04-30](#)



<1% match (student papers from 14-May-2017)  
[Submitted to Middle East University on 2017-05-14](#)



PADDY DISEASE DETECTION USING IMAGE PROCESSING BY ABID AL AMIN  
 ID: 151-15-5340 RAFEED AN NOOR ID:151-15-5364 AND MD. ROKIBUL  
 HASAN ID: 151-15-377 [This Report Presented in Partial Fulfillment of the  
 Requirements for the Degree of Bachelor of Science in Computer Science and  
 Engineering Supervised By Azizul Hakim Shuvo Lecturer Department of CSE  
 Daffodil International University DAFFODIL INTERNATIONAL UNIVERSITY  
 DHAKA, BANGLADESH APRIL 2019](#) [TABLE OF CONTENTS](#) [CONTENTS](#) [Board of  
 examiners Declaration Acknowledgements Abstract](#) [CHAPTER 1:  
 INTRODUCTION](#) [1.1 Introduction](#) [1.2 Background](#) [1.3 Problem Statement](#) [1.4  
 Expected Output](#) [1.5 Scope of Study](#) [1.6 Thesis Organization](#) [CHAPTER 2:  
 LITERATURE REVIEW](#) [2.1 Introduction](#) [2.2 Paddy Overviews](#) [2.1 Definition  
 of Paddy](#) [2.2.2 Paddy Disease Symptoms](#) [2.3 Case Study on Existing System](#)  
[2.4 Conclusion](#) [CHAPTER 3: PROPOSED METHODOLOGY](#) [PAGE i ii iii iv 1-3 1 1  
 2 3 3 3 4-7 4 4 4 5 6-7 7 8-20 i](#) [3.1 Introduction](#) [3.2 Workflow](#) [3.3 Image  
 Collection](#) [3.4 Image Processing](#) [3.5 Feature Extraction](#) [3.5.1 CCM](#) [3.5.2  
 Textural Features](#) [3.5.3 Feature Selection](#) [3.6 Classification](#) [3.6.1 Leaf Color  
 Analysis](#) [3.6.2 Artificial Neural Network](#) [CHAPTER 4: EXPERIMENTAL  
 ANALYSIS](#) [4.1 Introduction](#) [4.2 Experimental Setup](#) [4.3 Performance  
 Measures – Definition](#) [4.4 Result Analysis](#) [4.4.1 Random Test\(k-fold\)](#) [4.4.2  
 Classification Result](#) [4.5 Conclusion](#) [CHAPTER 5: CONCLUSION](#) [5.1 Summary](#)  
[5.2 Limitation](#) [5.3 Future Scope of Work](#) [REFERENCES](#) ©Daffodil International  
 University 8 8 8-9 9-10 10 11-14 14-17 18 18 19 19-20 21-27 21 21 22-23  
 23 24 25-27 27 28-29 28 28 29 30-32 ii [APPENDIX](#) [Appendix A: Research  
 Reflection PLAGIARISM REPORT](#) 33 33 34 iii [LIST OF FIGURES](#) [FIGURES](#) [PAGE  
 NO](#) [Figure 1.1: Rice Production and Consumption Statistics in Bangladesh](#) 2  
[Figure 3.1: Work Flow of proposed methodology](#) 8 [Figure 3.2: Sample of  
 collected images](#) 9 [Figure 3.3: L\\*a\\*b\\* Color Space in Dimensional Graph](#) 9  
[Figure 3.4: Processing image from RGB to Lab](#) 10 [Figure 3.5: Brown Spot  
 affected paddy image](#) 13 [Figure 3.6: Proposed Artificial Neural Network](#) 20  
[Figure 4.1: Correctly classified 4 Paddy Blast images](#) 25 [Figure 4.2:  
 Incorrectly classified 4 Paddy Blast images](#) 25 [Figure 4.3: Correctly classified](#)