Krishokbondhu - An automated system for diagnosis of paddy disease

Supervisor
Dr Mohammad Zahidur Rahman

Co-Super Visor
Md. Zahangir Alom

Conducted by
Shovon Paulinus Rozario
ID: 11101074

SCHOOL OF ENGINEERING AND COMPUTER SCIENCE
Department of Computer Science and Engineering
BRAC University

Submitted on September 1, 2014

DECLARATION

This is to certify that this thesis report is submitted by Shovon Paulinus Rozario (ID: 11101074) for the degree of Bachelor of Science in Computer Science and Engineering to the Department of Computer Science and Engineering, School of Engineering and Computer Science, BRAC University. I hereby declare that this thesis is based on the results found by myself and the materials of work found by other researchers are mentioned by reference. The contents of this thesis, neither in whole nor in part have been previously submitted to any other Institute or University for any degree.

Signature of Supervisor	Signature of Author
Dr Mohammad Zahidur Rahman	Shovon Paulinus Rozario

Signature of Co-supervisor
Md. Zahangir Alom

ACKNOWLEDGMENTS

I thank all who in one way or another contributed in the completion of this thesis. I am very much grateful to my supervisor Dr Mohammad Zahidur Rahman, Professor and Chairperson, BRAC University and co-supervisor Md. Zahangir Alom for guiding me developing this system regardless of their busy schedule. It would have been really difficult to bring this work towards a completion without their guidance, enormous encouragement and continuous support. I also want to thank Dr. Mohammad Ashik Iqbal Khan, Senior Scientific Officer, Bangladesh Rice Research Institute (BRRI), for his assistance regarding validation of the approach for the developed system and suggestions for improving this system furthermore in future. I am thankful to Md. Rezwanul Hoque as well for his efforts towards developing this system.

ABSTRACT

This paper demonstrates 'Krishokbondhu', an automated system for the farmers to identify paddy diseases using their mobile phones. The uploaded pictures captured by the mobile phones are processed in the remote server and presented to an expert group for their opinion. Computer vision techniques are used for detection of affected spots from the image and their classification. A simple color difference based approach is followed for segmentation of the disease affected lesions. Blob detection algorithm is used for feature extraction from the segmented images where features like number of blobs in the crop, nitrogen level of the leaf, area and color values of the affected areas etc are used for classification of the diseases. Binary Search Tree is used for mapping the feature values for comparison of Euclidean distance during classification. The system allows the expert to evaluate the analysis results and provide feedbacks to the famers through a notification to their mobile phones. The mobile application has been developed for Windows Phone and the remote server application is developed using .NET framework.

keywords: crop disease management system, delta E, blob detection, Euclidean distance, binary search algorithm

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER ONE

INTRODUCTION

Paddy is one of the most important crops all over the world considering its impact in the global food market. With the growing population all over the world, the demand for food item like rice is increasing more than ever. The environmental influences (i.e. soil, weather) on the cultivation of paddy have a major contribution to the production rate of rice all over the world. However, the next major influence on increasing production is the effective management of paddy diseases and pests. Farmers lose an estimated average of 37% of their rice crop to pests and diseases every year [1]. It is important that the farmers get to identify the condition of their paddy well ahead of time before it is too late, in order to avoid any kind disaster that can be caused by the diseases. Accurate diagnosis and timely solving of paddy disease is thus a vital component of rice production management aiming for enhanced productivity leading to increased profits.

This paper presents 'Krishokbondhu', an automated system integrated with machine vision techniques that will assist the farmers get the accurate information about their crops using their mobile phone. The uploaded pictures of paddy captured by the mobile phones will be processed in the central server and the analysis report will be presented to an expert group for their opinion, who will then be able to send proper recommendations through a simple notification using the system, according to the severity of the situation.

1.1 Motivation

In today's digital age, it is important that the farmers get to use the latest technology for efficient management of their crops. The use of information access through mobile phones among the farmers has increased in recent years, which has made a positive impact on the output of the production. However, there is still a lacking of knowledge sharing between the farmers and the agriculture experts while it comes on a topic of proper crop management due to the challenges in training the farmers on a mass level on topics like disease identification and their management. As a result, in most of the cases the farmers rely on their experience and intuition for decision on identifying crop diseases and their treatments. The production might turn out not as expected if the symptoms are not treated in a proper manner, using appropriate amount of fertilizers guided by agriculture specialist. The motivation of this paper is to provide the farmers the access to a service which will connect them directly to the specialist for serving their needs on effective management of disease. As a first step, this system focuses on creating a service for management of paddy diseases using mobile phones. Image processing techniques have been applied in this study for identification of four paddy diseases named Bacterial Leaf Blight, Brown Spot, Leaf Blast and Leaf Scald. It is possible to make this an autonomous system for disease identification and providing suggestions based on image analysis reports, but it is essential to govern the data by an expert so that the farmers get the best possible solution for their problems regarding paddy cultivation.

1.2 Thesis Outline

Chapter-2 reviews the background study for development of the current system, which includes the detailed aspects of the targeted paddy diseases followed by the literature review for related works in this field.

Chapter-3 presents the system design of the developed system. It includes various figures and diagrams that were used for designing and developing 'Krishokbondhu'.

Chapter-4 describes the implementation details of the application, including the analysis techniques used for image processing based disease recognition approach.

Chapter-5 reviews the results of the research and provides a discussion on the project findings.

Chapter-6 specifies the limitations of the system and provides the conclusion and future guidelines.

CHAPTER TWO

BACKGROUND STUDY

2.1 Paddy Diseases

Disease damage to rice can greatly reduce yield. They are mainly caused by bacteria, viruses, or fungi [3]. In most of the cases the diseases create visual symptoms, primarily creating spots or changing color on the leaf body, tip or stem of paddy. The most common diseases of paddy are Bacterial Leaf Blight, Brown Spot, Leaf Blast, Leaf Scald, Narrow Brown Spot, Bacterial Leaf Streak, False Smut, Sheath Blight, Red Stripe, Stem Rot etc [3]. For applying machine vision based disease recognition based on visual symptoms, this paper focuses on four diseases named Bacterial Leaf Blight, Brown Spot, Leaf Blast and Leaf Scald.

Bacterial Leaf Blight

Bacterial Leaf Blight is caused by *Xanthomonas oryzae* pv. oryzae. It causes wilting of seedlings and yellowing and drying of leaves. The disease is most likely to develop in areas that have weeds and stubbles of infected plants. It can occur in both tropical and temperate environments, particularly in irrigated and rainfed lowland areas.

Figure 2.1: Bacterial Leaf Blight

On seedlings, infected leaves turn grayish green and roll up. As the disease progresses, the leaves turn yellow to straw-colored and wilt, leading whole seedlings to dry up and die [3].

Brown Spot

Brown Spot is caused by a fungus called *Cochliobolus miyabeanus*. Brown spot has been historically largely ignored as one of the most common and most damaging rice diseases. Brown spot is a fungal disease that infects the cleoptile, leaves, leaf sheath, panicle branches, glumes, and spikelets. Its most observable damage is the numerous big spots on the leaves which can kill the whole leaf. When infection occurs in the seed, unfilled grains or spotted or discolored seeds are formed, especially with round to oval, dark-brown lesions with yellow or gold halo. As lesions enlarge, they remain round, with center area necrotic gray and the lesion margin reddish-brown to dark brown [3].

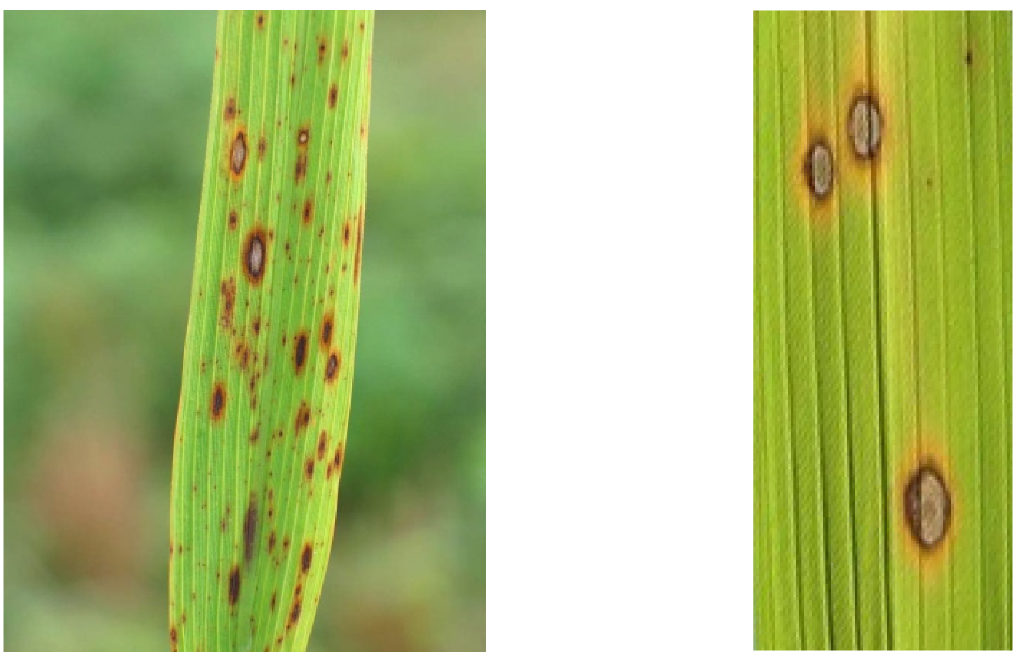


Figure 2.2: Brown Spot

Leaf Blast

Blast is caused by the fungus *Magnaporthe oryzae*. It can affect all above ground parts of a rice plant: leaf, collar, node, neck, parts of panicle, and sometimes leaf sheath. Blast can occur wherever blast spores are present. It occurs in areas with low soil moisture, frequent and prolonged periods of rain shower, and cool temperature in the daytime. In upland rice, large day-night temperature differences that cause dew formation on leaves and overall cooler temperatures favor the development of the disease. Initial symptoms appear as white to gray-green lesions or spots, with dark green borders. Older lesions on the leaves are elliptical or spindle-shaped and whitish to gray centers with red to brownish or necrotic border. Some resemble diamond shape, wide in the center and pointed toward either end. Lesions can enlarge and coalesce, growing together, to kill the entire leaves [3].

Figure 2.3: Leaf Blast

Leaf Scald

Leaf scald is a fungal disease caused by *Microdochium oryzae,* which causes the scalded appearance of leaves. Disease development usually occurs late in the season on mature leaves and is favored by wet weather, high nitrogen fertilization, and close spacing. It develops faster in wounded than unwounded leaves. Individual lesions are 1−5 cm long and 0.5−1 cm wide or may almost cover the entire leaf. The continuous enlargement and coalescing of lesions result in blighting of a large part of the leaf blade. The affected areas dry out giving the leaf a scalded appearance [3].

Figure 2.4: Leaf Scald

2.2 Literature Review

The application of machine vision techniques have increased broadly in the agriculture industry in last few years, especially in the plant protection field which ultimately leads to crops management. However, there are not many applications that offer portable solution that enables management of paddy using image analysis techniques. Amos Gichamba and Ismail Ateya Lukandu in [20] described different implementations of mobile systems in agriculture and presented a model for designing such applications. The paper finds that the development of mobile solutions in the agriculture sector has not been done widely and shows how solutions can be created using mobile technology that will help in addressing some of the problems faced by farmers [20]. International Rice Research Institute (IRRI) developed a web based diagnostics tool named Rice doctor for identifying problems in crop in order to provide actionable advice how to manage them [2], where the suggestions are generated based on manual selection of information provided by the farmers. Phadikar, S. and Sil, J. presented a prototype system for rice disease detection based on the infected images of various rice plants [16]. Images of the infected rice plants are captured in the system by digital camera and processed using image growing, image segmentation techniques to detect infected parts of the plants. Then the infected part of the leaf has been used for the classification purpose using neural network [16]. Reinald Adrian D. L. Pugoy and Vladimir Y. Mariano developed a system where the outlier region is first obtained from a rice leaf image to be tested using histogram intersection between the test and healthy rice leaf images [17]. Upon obtaining the outlier, it is then subjected to a threshold-based K-means clustering algorithm to group related regions into clusters. Then, these clusters are subjected to further analysis to finally determine the suspected diseases of the rice leaf [17]. Among the mobile application based systems, [18] presented 'Beetle' which is designed to support farmers in rural area to detect crop disease. 'Beetle' detects crop disease from the image captured by a cell phone and detects the disease in real time using histogram and color information of the image. Similar approach has been proposed by [19] to detect the shapes of medicinal plants.

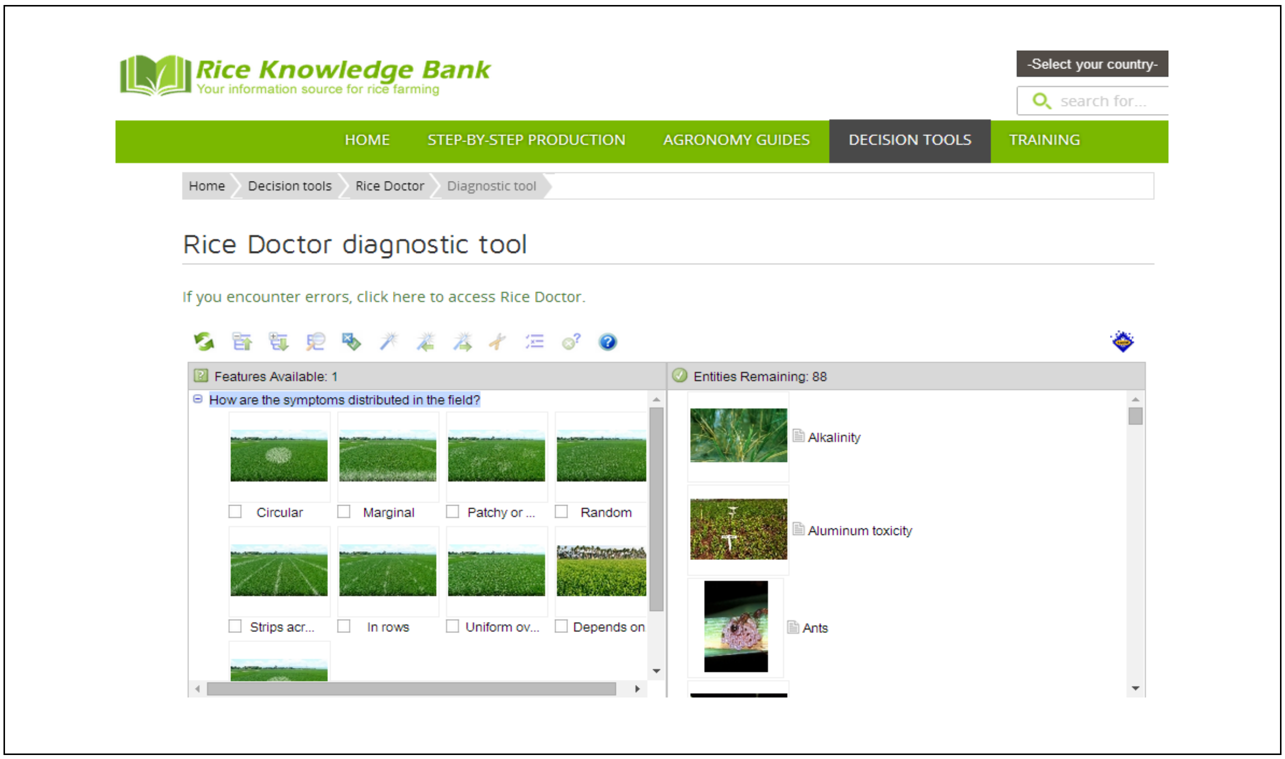


Figure 2.5: Rice Doctor paddy diagnostic tool

CHAPTER THREE

SYSTEM MODEL

3.1 System Overview

The system consists of a mobile application, which will enable the farmers to take images of paddy using their mobile phones and send it to a central server where the central system in the server will analyze the pictures based on visual symptoms using image processing algorithms in order to measure the disease type. An expert group will be available to check the status of the image analysis data and provide suggestions based on the report and their knowledge, which will be sent to the farmer as a notification in the application.

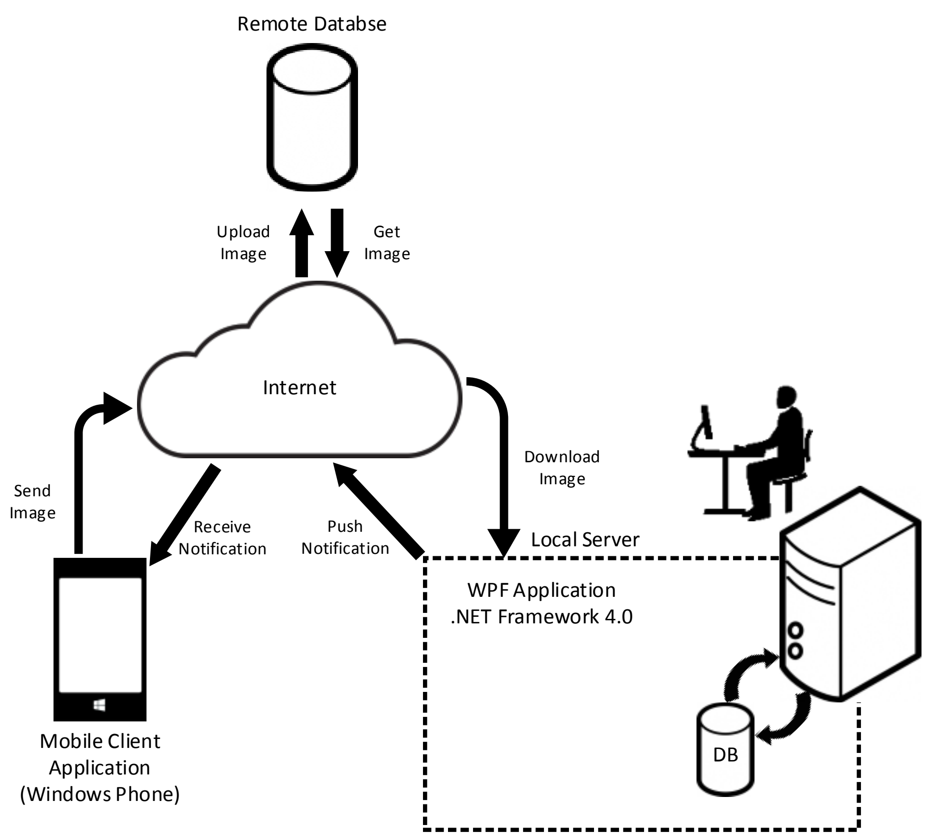


Figure 3.1: General system architecture

3.2 Activity Diagram

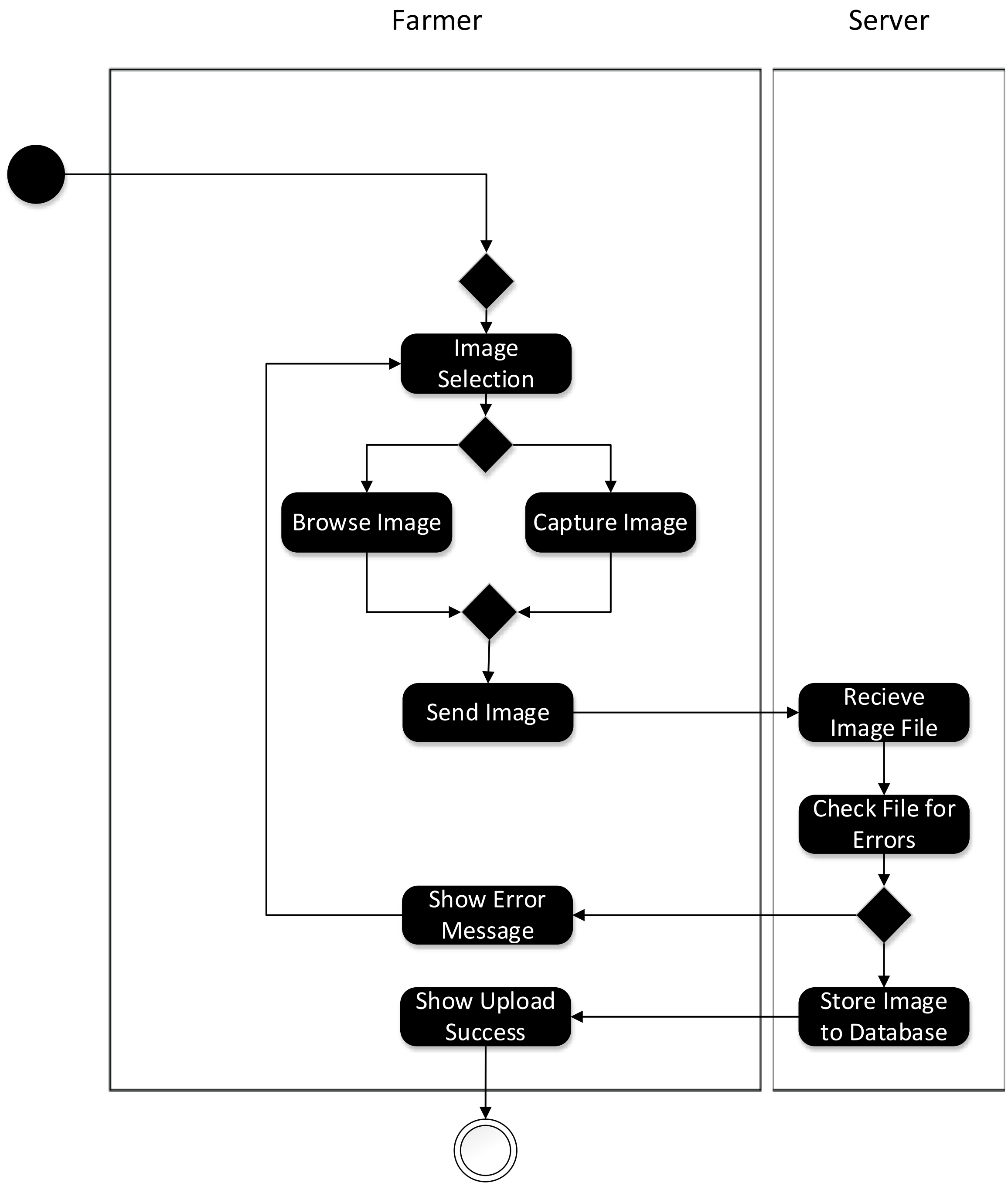


Figure 3.2: Image upload activity diagram

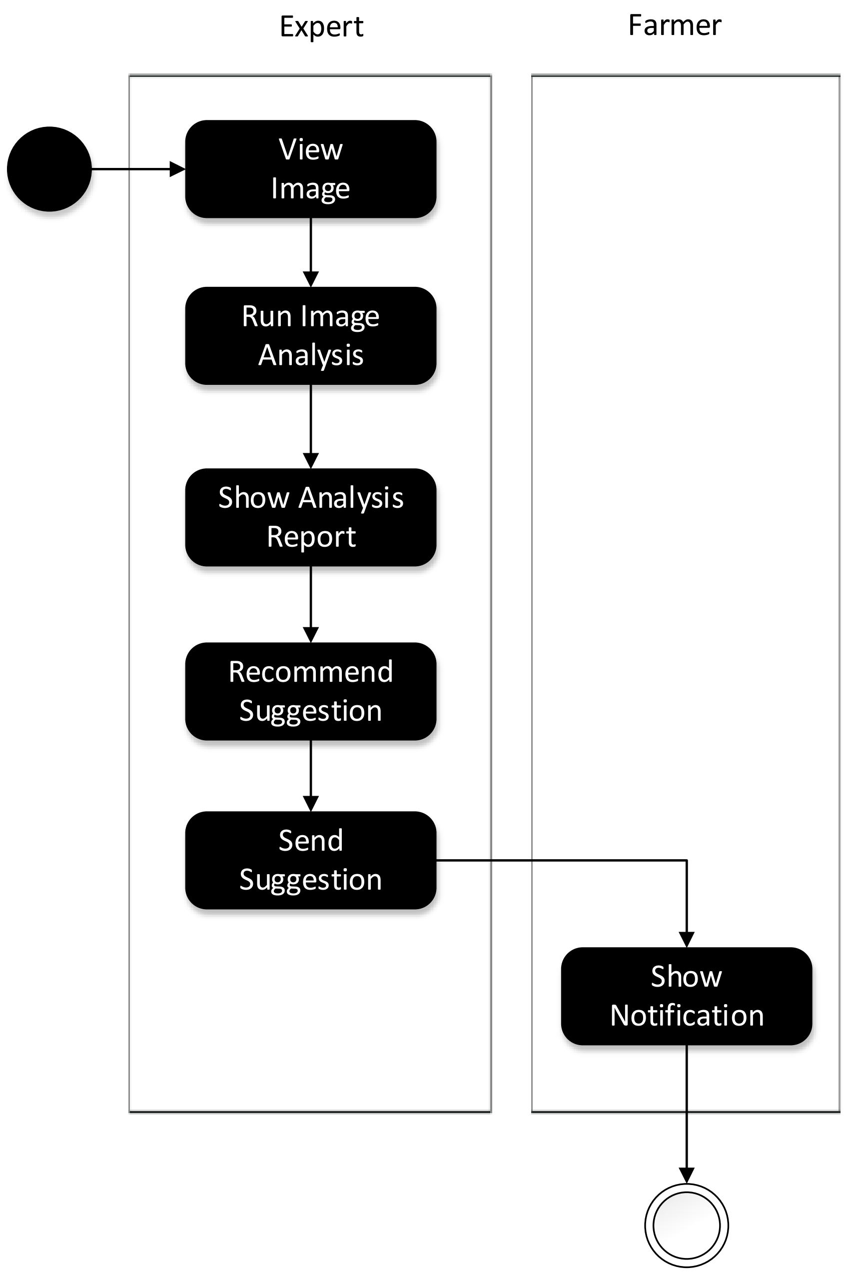


Figure 3.3: Expert analysis and feedback activity diagram

3.3 Use Case Diagram

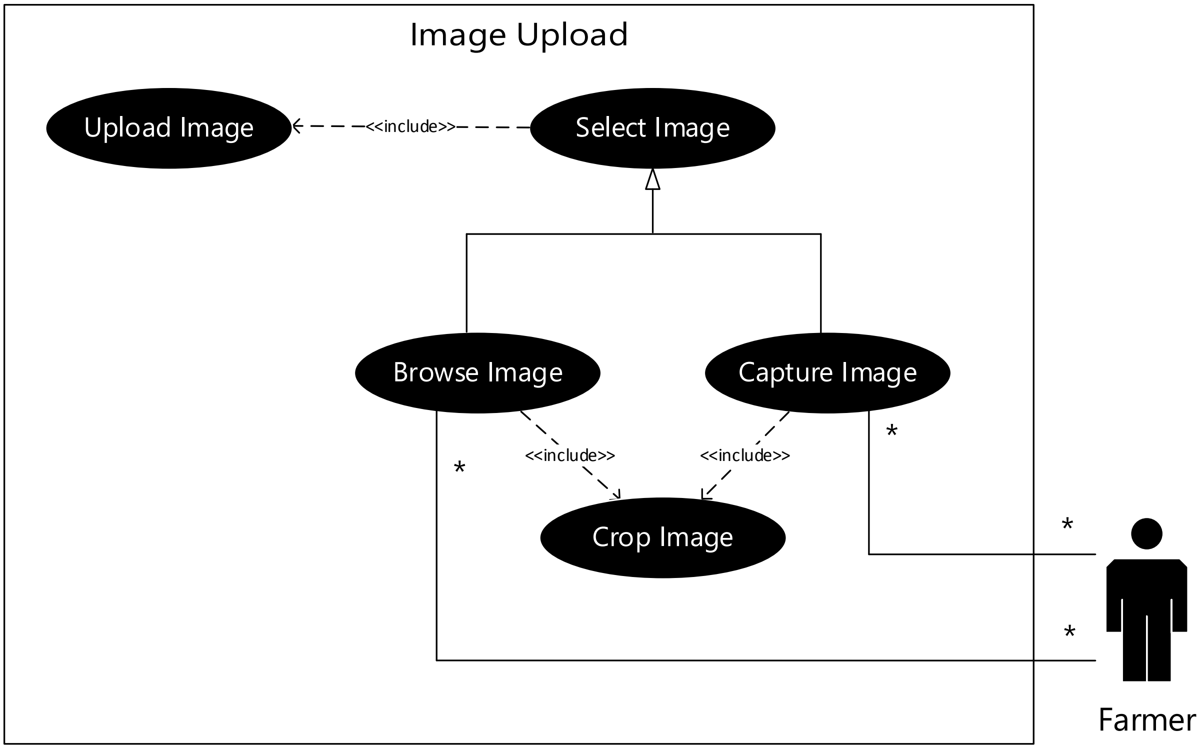


Figure 3.4: Image upload use case diagram

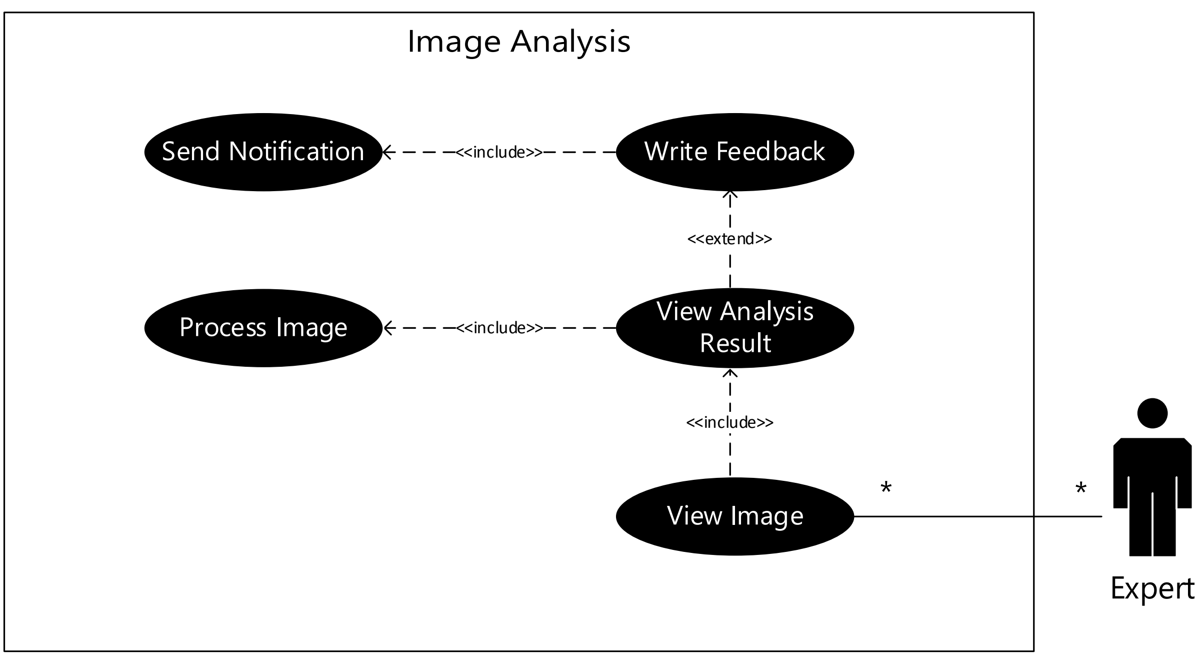


Figure 3.5: Image analysis use case diagram

3.4 Sequence Diagram

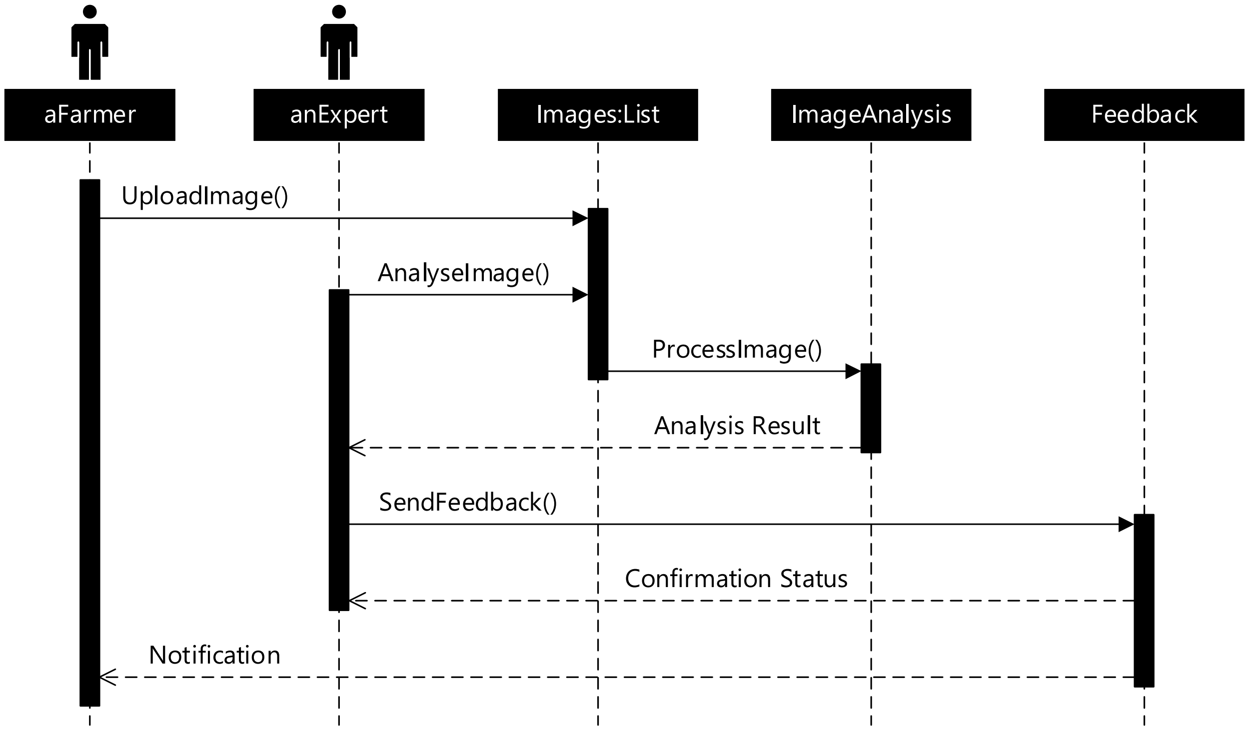


Figure 3.6: Sequence diagram

3.5 Class Diagram

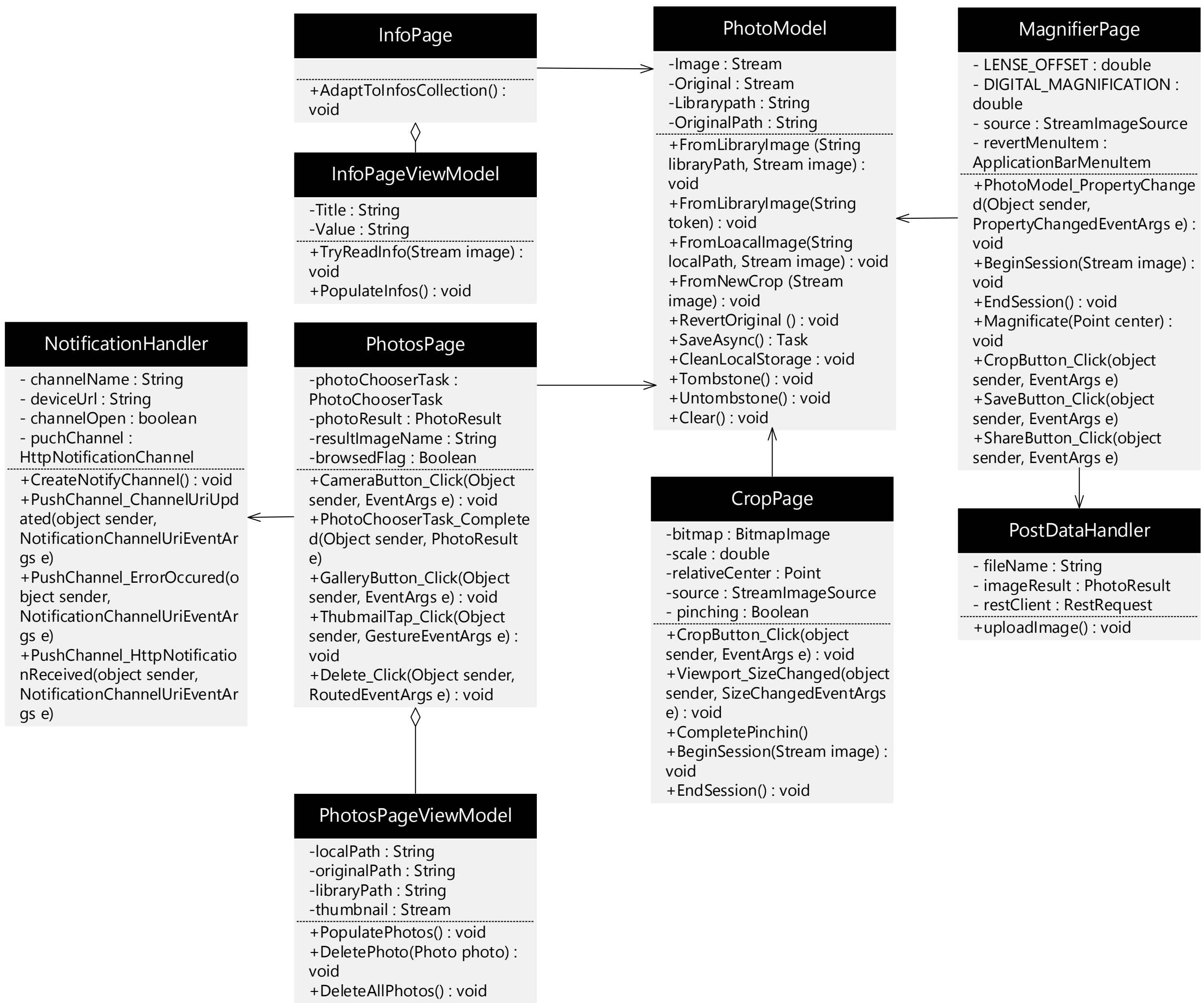


Figure 3.7: Windows phone application class diagram

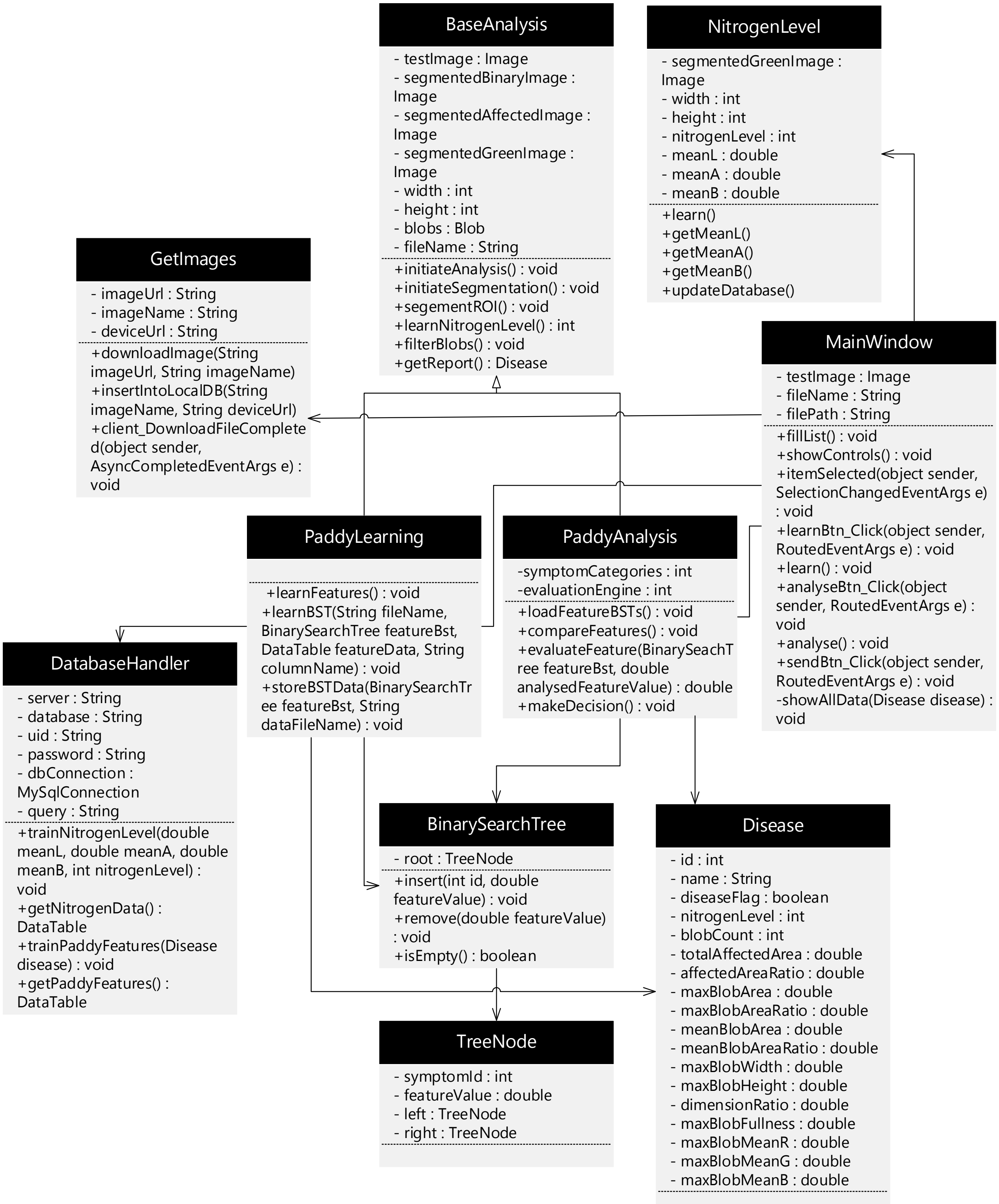


Figure 3.8: WPF server application class diagram

CHAPTER FOUR

SYSTEM IMPLEMENTATION

4.1 Mobile Application Development

The mobile application for the system 'Krishokbondhu' is targeted to be used by the farmers. Nokia being part of Windows in recent time, it will be within the reach of the farmers to get access to a low price smart phone which they can use for using the application 'Krishokbondhu'. Thus the client mobile application has been developed for windows phone 8, using C# as programming language. Nokia Imaging SDK was used for implementation of the basic imaging functionalities with the application.

The mobile application consists of 5 basic functionalities. They are 1) Image capture, 2) Image selection, 3) Image zoom and crop, 4) Share image with expert group, 5) Receive notification from central server.

1) Image capture

At the very first page of the application, the application bar shows the icon for capturing

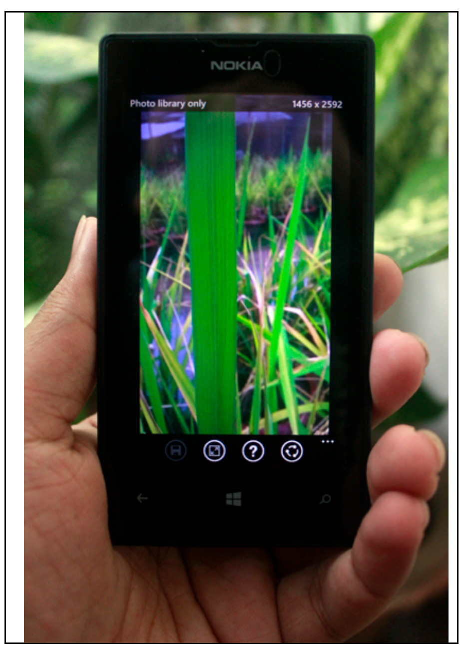


Figure 4.1: Photo capture using mobile application

image using the application. On navigation of the menu, the user gets to take image on shutter click event using the phone.

2) Image selection

In case of previously taken pictures of paddy, the application navigation menu also contains the option of selecting an image from the existing photo library of the phone.

3) Image zoom and crop

The leaf of paddy is a very thin one, and it is important that the targeted area of the leaf gets focus in the image. The mobile application of 'Krishokbondhu' lets the farmer to zoom the affected region of paddy using pinch with two fingers. The test images were taken with a Nokia N520 phone which has a 5 mega pixel camera in it. The application allows to zoom 4x times the original image. In addition, once the targeted region has been selected, the crop button of the crops the image in a 170x400 pixel frame, which is

Figure 4.2: Selecting target region of paddy leaf in mobile application

the targeted resolution for processing images in the server image processing application.

4) Share image with expert group

This functionality of the application enables the farmer to send the captured/ selected image to be uploaded in the remote server using HTTP Web Services in Windows phone. The client mobile application uploads two basic types of data in the server for every request, the image that the farmer selects for seeking suggestion and a unique URL created through windows live services which is created for communicating with the mobile phone from a remote application. The URL created using Windows Live service is used for sending notification from the server application of ‘Krishokbondhu’, sent by the expert groups with their feedback. For uploading the data in the remote server as a POST type data, ‘RestSharp’ which is a REST client library, has been used. The method for upload is selected as POST, along with providing the url to the remote server script for receiving the file while creating a request using RestRequest. The request then adds the parameter (device url generate by Windows Live) and the file to it and is sent to the server through a object of RestClient which uses Asynchronous Task execution for uploading the post data.

```
//preparing RestRequest by adding server url, parameteres and files...
RestRequest request = new RestRequest("http://krishokbondhu.leninhasda.net/
  upload.php", Method.POST);
request.AddParameter("deviceUrl", NotificationHandler.deviceUrl);
request.AddFile("photo", ReadToEnd(PhotosPage.resultImage),
  PhotosPage.resultImageName, "image/pjpeg");

//calling server with restClient
RestClient restClient = new RestClient();
restClient.ExecuteAsync(request, (response) =>
{
    if (response.StatusCode == HttpStatusCode.OK)
    {
        //upload successfull
        MessageBox.Show(response.Content);
    }
    else
    {
        //error ocured during upload
        MessageBox.Show(response.StatusCode + "\n" + response.StatusDescription);
    }
});
```

Figure 4.3: Code snippet for POST image and data using RestSharp

Asynchronous programming helps the app stay responsive when it does work that might take an extended amount of time [4]. For example, while uploading the contents to the server it might spend several seconds waiting for the content to upload. In case of using a synchronous method on the UI thread to upload the content, the application is blocked until the method returns [4]. The app won't respond to user interaction, and because it will seem non-responsive which might turn the user frustrated. Thus the asynchronous task is used, where the app continues to run and respond to the UI while it waits for an operation to complete. Upon completion of its operation, the response of the task will be showed to the client. If there is any problem with the connection or uploading of the data, it will also notify the user in that case.

5) Receive notification from central server

Once the image has been uploaded in the remote server, the expert sends feedback to the client mobile application via notification. This notification is sent through a URL

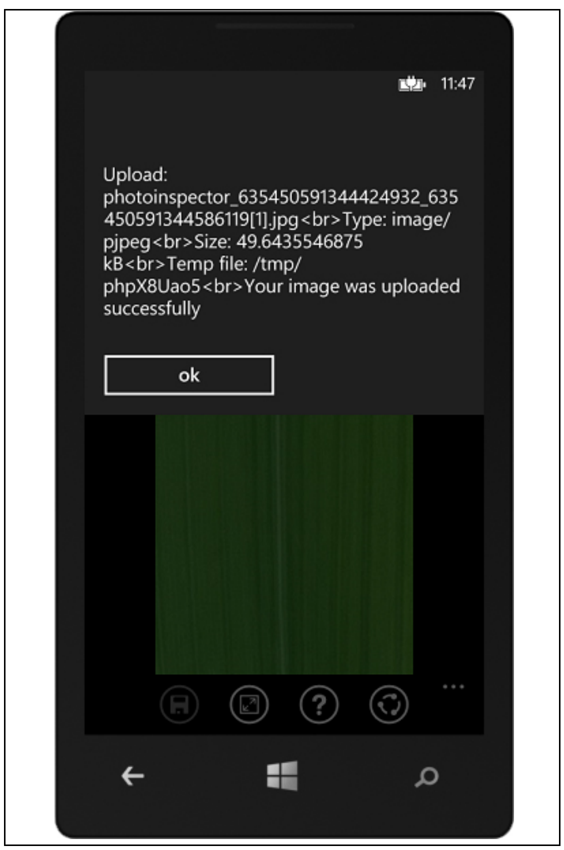


Figure 4.4: Image upload success confirmation

generated by Windows Live Services which is unique for every device. Once the notification is received, it is displayed in the application which the user will be able to view for taking appropriate steps suggested by the experts.

4.2 Remote Server Script for Storing Data

The image and the data uploaded from Windows Phone application is received in the remote server using a PHP script. The images are stored within the assigned directory and the device URL is mapped in the database against every uploaded image in the database.

4.3 Local Server Application Development

The local server application is dedicated for the use of the expert/ expert group. The server application is a Windows Presentation Form (WPF) application developed in Microsoft .NET framework 4.0 using C# programming language. The purpose of this application is to provide three functionalities to the expert, 1) Train paddy disease knowledgebase, 2) Analyze images sent by the farmers, and 3) Send suggestion to the farmers through notification.

Every image in the server application is analyzed using machine vision techniques for identification of possible disease of paddy presented in the image. Figure **CHECK** shows the steps of image processing for both training and analysis stages of the process cycle. The developed system works for detecting diseases for four types for the paddy, they are – Bacterial Leaf Blight, Brown Spot, Leaf Blast and Leaf Scald. For analysis of the images, Emgu CV and AForg.NET image processing library has been used. An image goes through the following steps for analysis.

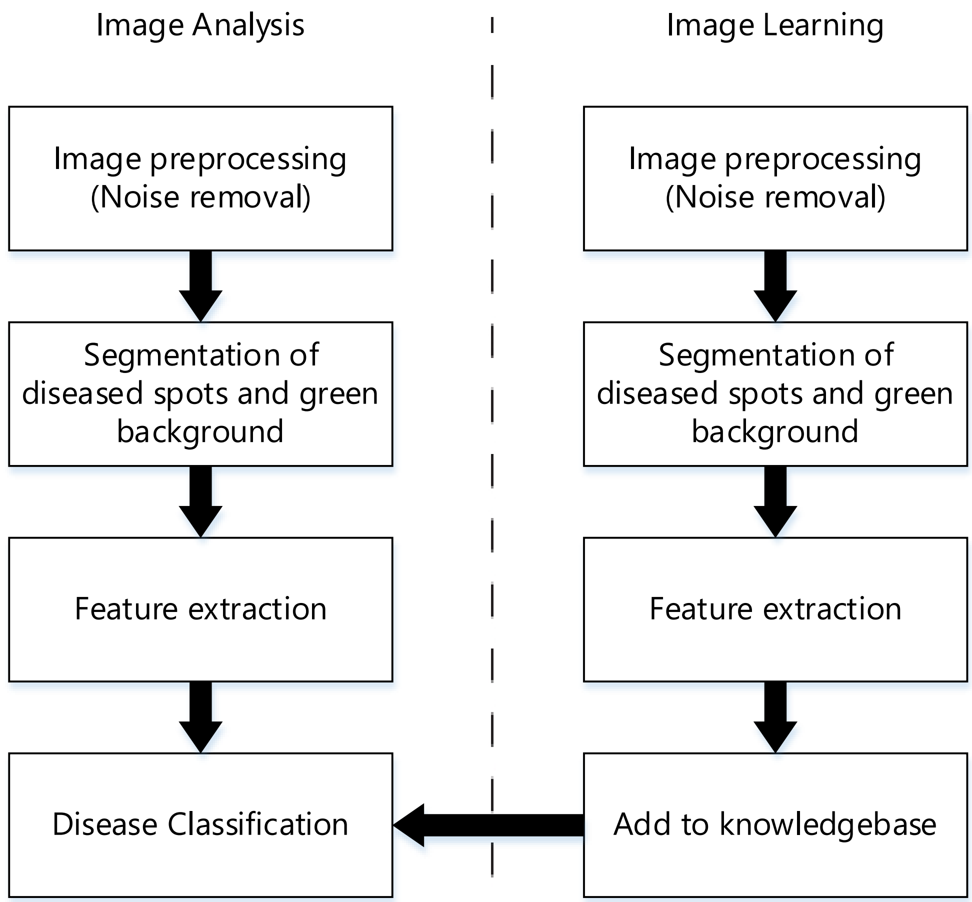


Figure 4.5: The basic procedures of image-processing for disease detection

4.3.1 Noise removal

The uploaded images may contain noise, for which a bilateral smoothing filter has been for noise cancellation. The bilateral filter is a technique to smooth images while preserving edges where the intensity value at each pixel in an image is replaced by a weighted average of intensity values from nearby pixels [6]. Bilateral filter is used for noise removal from medical images [6]. Emgu CV image processing library provides the functionality for application of this noise removal technique.

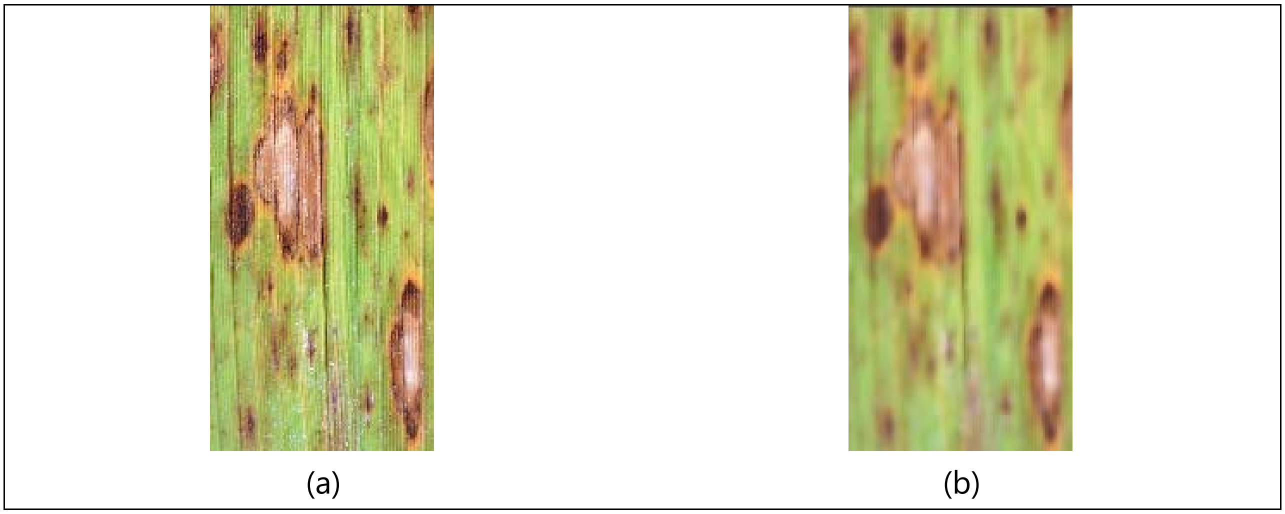


Figure 4.6: Before (a) and after (a) applying bilateral smoothing filter

4.3.2 Segmentation

A simple color feature based approach has been followed for segmentation of the disease affected images of the paddy. Color is an important feature in color image processing, especially in crop images. It has been observed for the paddy diseases that the RGB value of the affected region in a paddy can be key component for separating the target region from the non-affected one [7]. The non-affected paddy leaf is usually is green in color, leading the value of Green color to be higher than either of the Red or Blue color for each pixel. However, in case of the affected spots, the value of Green pixel

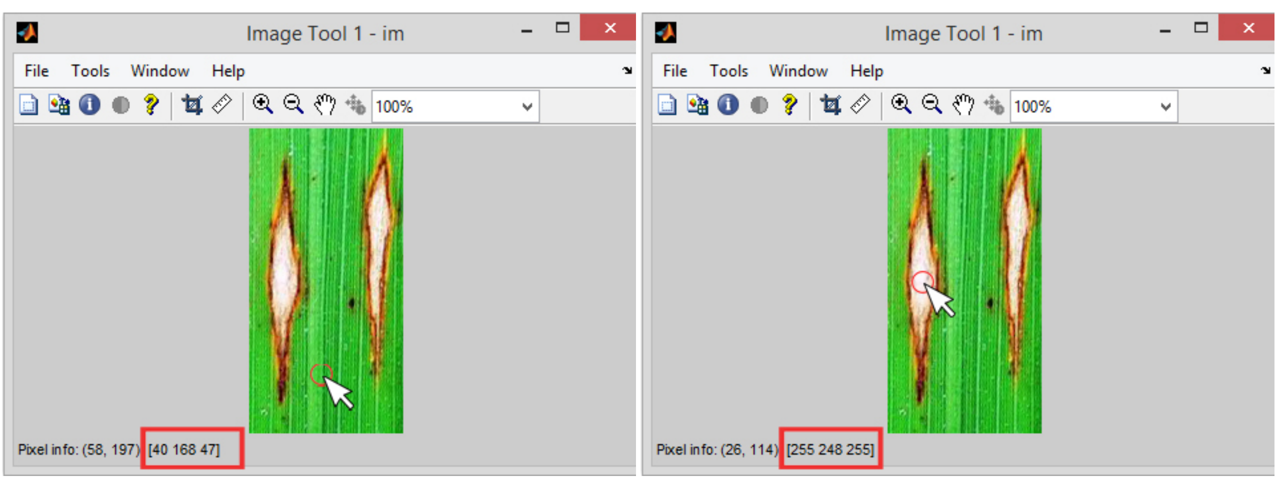


Figure 4.7: Color difference of RGB values between the affected region and green part of paddy

is less than either of the Red or Blue values. So after initial recognition of the green region within the image, we find mean RGB color value of the green part and convert it lab color space for finding delta E. Delta E is the color difference - between a pixel and its neighbors [6]. Once the mean LAB color value of the green area has been calculated, the distance delta - E (ΔE_{ab}) is calculated for every pixel of the image using the following equation,

$$\Delta E_{ab} = \sqrt{((L_m - L_i)^2 + (A_m - A_i)^2) + (B_m - B_i)^2)}$$

Here, L_m = mean L value; A_m = mean A value; B_m = mean B value; L_i = L value for ith pixel; A_i = A value for ith pixel; B_i = B value for ith pixel. ΔE_{ab} represents the distance of the pixel value from the mean green region pixel value and the non-green pixel value will be at far more distance [8], whereas the green pixels will tend to be close to the mean value and the ΔE_{ab} will be minimal (close to 0). The maximum color difference within the green region and the mean pixel value is set the threshold value for segmentation [15]. In this process we segment the diseased paddy image and remove the pixels for which the value of ΔE_{ab} is greater than the calculated threshold value.

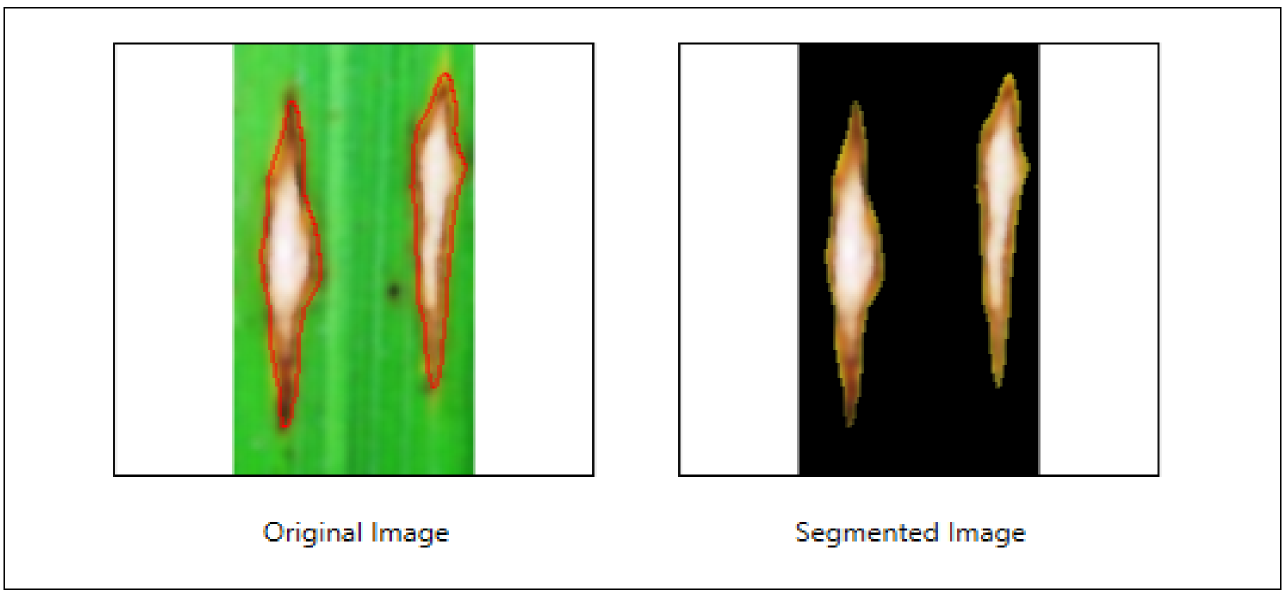
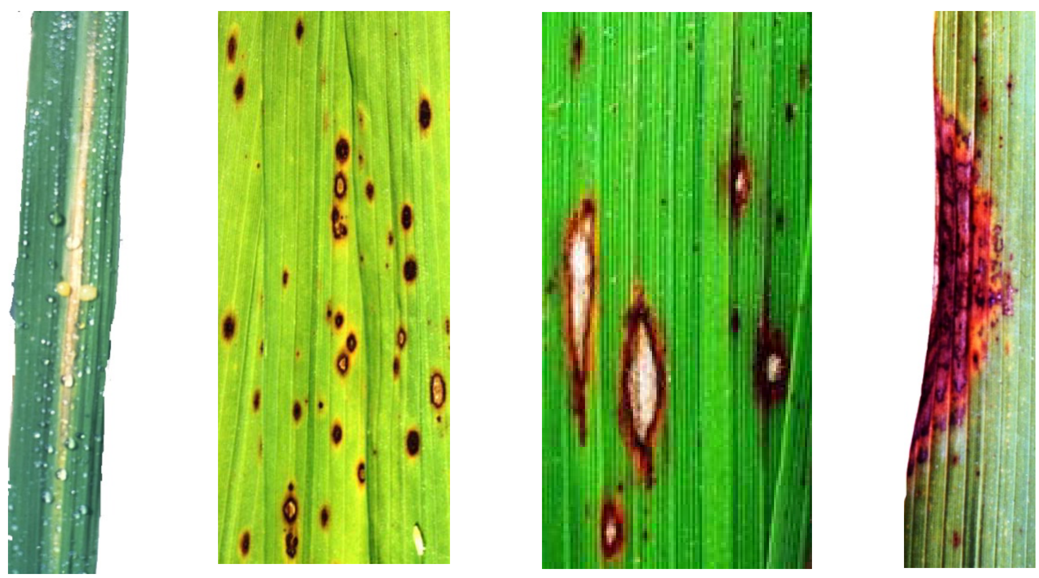


Figure 4.8: Original and segmented paddy leaf image

4.3.3 Feature Extraction

The goal of the feature extraction is to characterize an object to be recognized by measurements whose values are very similar for objects in the same category, and different for objects in other categories [9]. For the selected diseases that the designed system analyzes, the nature of visual symptoms provides the basis for which the features can be generated.



(a) (b) (c) (d)

Figure 4.9: Variety in visual symptoms for different paddy diseases, (a) Bacterial leaf blight, (b) Brown spot, (c) Leaf blast, (d) Leaf scald

The distinctions in the pattern of the affected area in the diseased images are visible. It is noticeable that the lesions/ damaged part of the leaf are the key component of creating variety among different diseases. There are also other observations like color difference among the non-affected regions of the leaves. A survey (Appendix A) was conducted for validating the features of the targeted paddy diseases by Plant Pathology experts of Bangladesh Rice Research Institute (BRRI). Based on the recommendations, the following visual features have been evaluated for comparison of paddy diseases. The spores of the diseased leaf has referred as blob.

- Nitrogen level
- Number of lesions
- Color for centre of the lesion
- Color for boundary of the lesion
- Area of diseased region
- Ratio of affected region (with the entire area)
- Average lesion area
- Biggest lesion area
- Ratio of the biggest lesion area
- Biggest lesion height
- Biggest lesion width
- Dimension ratio of the biggest lesion
- Fullness of the biggest lesion
- Average color value of the lesions

Measuring nitrogen level of the paddy leaf is the first step of feature extraction in the systems. The interaction between mineral nutrients like nitrogen and foliar pathogens is an important factor affecting the growth and production of field crops [10]. It may either increase or decrease the resistance or the tolerance of plants to pathogens by effecting changes in growth pattern, plant morphology and particularly chemical composition [10]. Greater accumulation of N has been reported to be responsible for lowering the silicated epidermal cell, decreasing hemicelluloses and lignin content (Matsuyama 1975) in the host tissue, thereby reducing the level of resistance to the disease [11]. On the other hand, soils fertilized with both high and low nitrogen levels can also increase rice susceptibility to brown spot [12]. Generally the Leaf Color Chart (LCC), jointly developed by International Rice Research Institute (IRRI) and Philippines Rice Research Institute is used to estimate the rice crop's nitrogen needs [13].

For measuring the nitrogen level of the paddy leaf from image, a Leaf Color Chart was used to collect sample data of the different N- levels in 9 different intensities of light. It is important to store data in different light intensity due to the fact that the color of paddy might differ for different weather conditions, which might lead to false conclusion

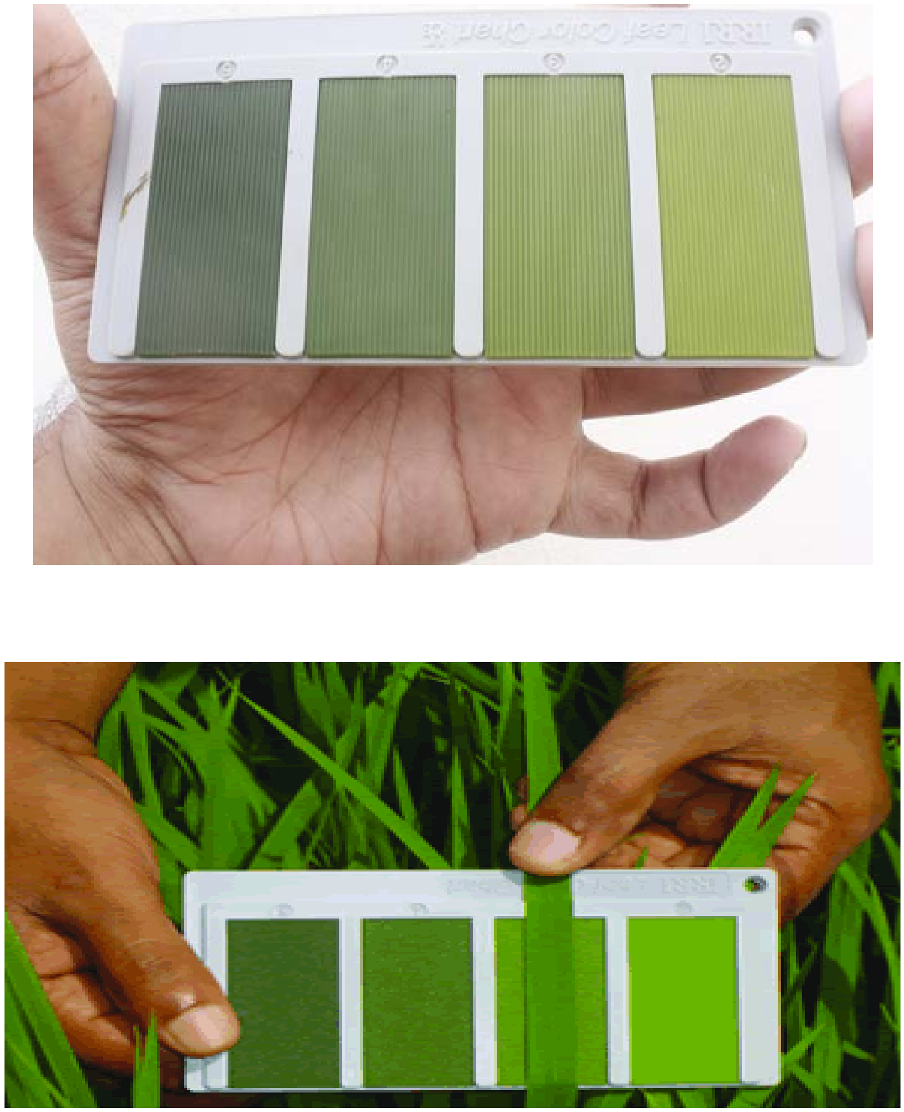


Figure 4.10: Measuring N- level using Leaf Color Chart (LCC)

of N- level for the leaf. The mean LAB color value for each sample image of the different N- levels were then calculated and stored in the database. In the feature extraction stage, while segmenting the diseased spots from paddy leaf image, the mean LAB value of only the green regions of the image is measured. The distance between the obtained mean value and the sample LCC values were then measured using the delta E equation and the one with the nearest distance represented the Nitrogen level for the diseased leaf.

$$\Delta E_{ab} = \sqrt{((L_m - L_i)^2 + (A_m - A_i)^2) + (B_m - B_i)^2)}$$

Maximum of the other key features are collected from the affected lesions of the diseased paddy leaf. Blob Detection algorithm has been applied for extracting these features from the disease affected spots, which can referred to blob as well. In the field of computer vision, blob detection refers to mathematical methods that are aimed at detecting regions in a digital image that differ in properties, such as brightness or color, compared to areas surrounding those regions [4]. AForge.NET image processing Library has been used for applying the blob detection algorithm in order to extract the features of the lesions from the segmented image of paddy. The Blob class of the library contains the properties like blob area, center of gravity, color mean, fullness, blob containing rectangle etc. [14]. The segmented image is thus processed using the given blob detection algorithm and the selected features of the blobs are analyzed. It is noticeable that many of the analyzed features contain properties for the largest lesion/ blob. It has been found from the survey (Appendix A) that the largest diseased spot of the leaf shows the extreme properties for the visual symptoms of the particular disease and thus is justifiable to consider properties of the largest blob. For example, in case of Leaf Blast disease the leaf may contain more than one spots in the leaf where the center of the lesion tends to be white in color, which might not visible in case of early stage of the spores.

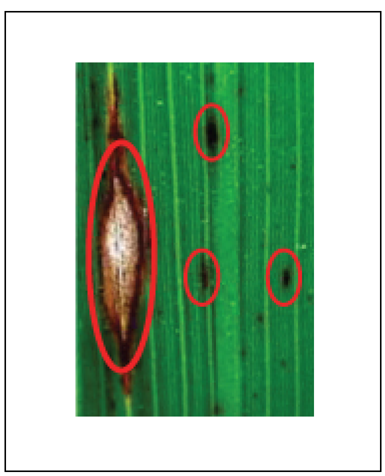


Figure 4.11: Lesions in the Leaf Blast disease

Thus, after filtering the blobs from the segmented image, the largest of the blobs is selected based on values of area of the lesions.

The two features, boundary color value and color for center of the lesion are not measured by the blob detection algorithm. We used seeded region growing algorithm for extracting these two features. Seeded region growing algorithm (SRG) is a approach which is based on conventional postulate of region growing algorithms where the criteria of similarity of pixels is applied [15]. The first step in region growing is to select a set of seed points and then grown from these seed points to adjacent points depending on a region membership criterions (like pixel intensity, color etc.). In the developed system, instead of selecting the center of the lesion as seed point, we first find the mean color value from boundary of the biggest lesion as seed point and grow from there for finding color value for center of the lesion. Shilpa Dantulwar and R. K. Krishna in [15] used delta E equation of RGB value for finding the similarity index between two neighbor pixels. The maximum color difference within the boundary region and the mean pixel value is set as the threshold value for evaluation of the similarity index [15].

$$\Delta E = \sqrt{}(D_r + D_g + D_b)$$

Where,

$D_r = \left(\left(R_{x+i,y+j} - R_{x,y}\right)^2\right)$

$D_g = \left(\left(G_{x+i,y+j} - G_{x,y}\right)^2\right)$

$B_r = \left(\left(B_{x+i,y+j} - B_{x,y}\right)^2\right)$

4.3.4 Add to knowledgebase

All the values of the features that have been considered are of numeric type. Since these values are used to construct the knowledgebase for comparison with the input image later on, it will be a time consuming process to go through every knowledge data one by one. We thus stored the feature values of each type using serialized Binary Search Tree object for application of Binary Search algorithm in the classification part.

4.3.5 Classification

We used the idea of maximum feature similarities as the basis for classification of the input image. In this stage, we measure the feature values of the input image and compare the Euclidean distance with features of already the learned images.

$$D_f = F_{input} - F_{learned}$$

Here,

D_f = Euclidean distance between feature of a input image and already learned image,

F_{input} = feature value for the input image

$F_{learned}$ = feature value for the learned image

Binary Search algorithm has been used for finding the lowest Euclidean distance of the feature. Using this algorithm, we find the learned value closest to the observed value using binary search, avoiding the comparison time for the values which are already at larger distance than the current one. This helps decreasing the search time to $log_2(n)$ in

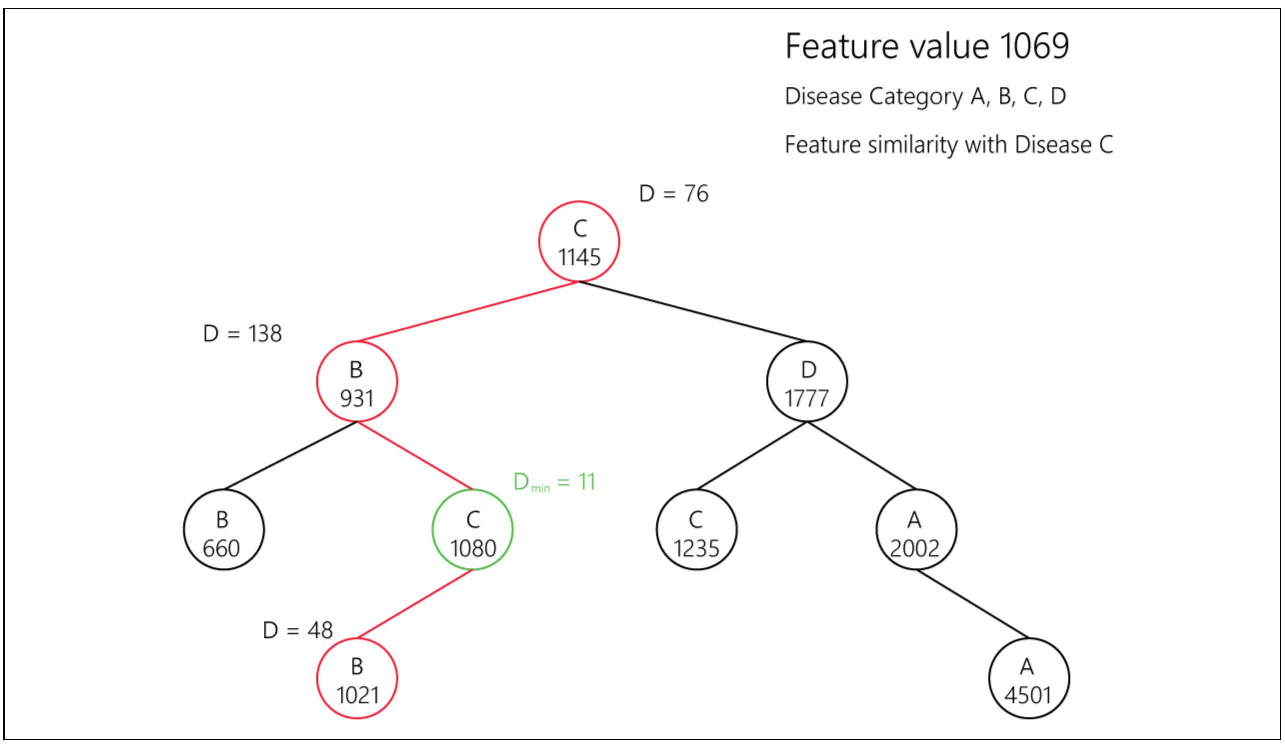


Figure 4.12: Comparison of Euclidean distance using Binary Search Algorithm

best case. However, this performance depends on the perfection of the binary tree where every node has exact two child nodes, which is a challenging task. In the worst where every node has only one child node, the time complexity will be linear of the learned samples (n). To minimize this worst case scenario, we select the feature value closest to the median value of that feature as the root node so that the tree becomes balanced.

Now for every feature, the disease with the minimum distance will increase the probability of the input image to be of that corresponding disease category. Euclidean distance for all the features are measured in this way, and the probability values are recorded for the corresponding diseases. The disease category with the most number of feature similarities is marked as the diseased category for the input image and is being displayed in the output.

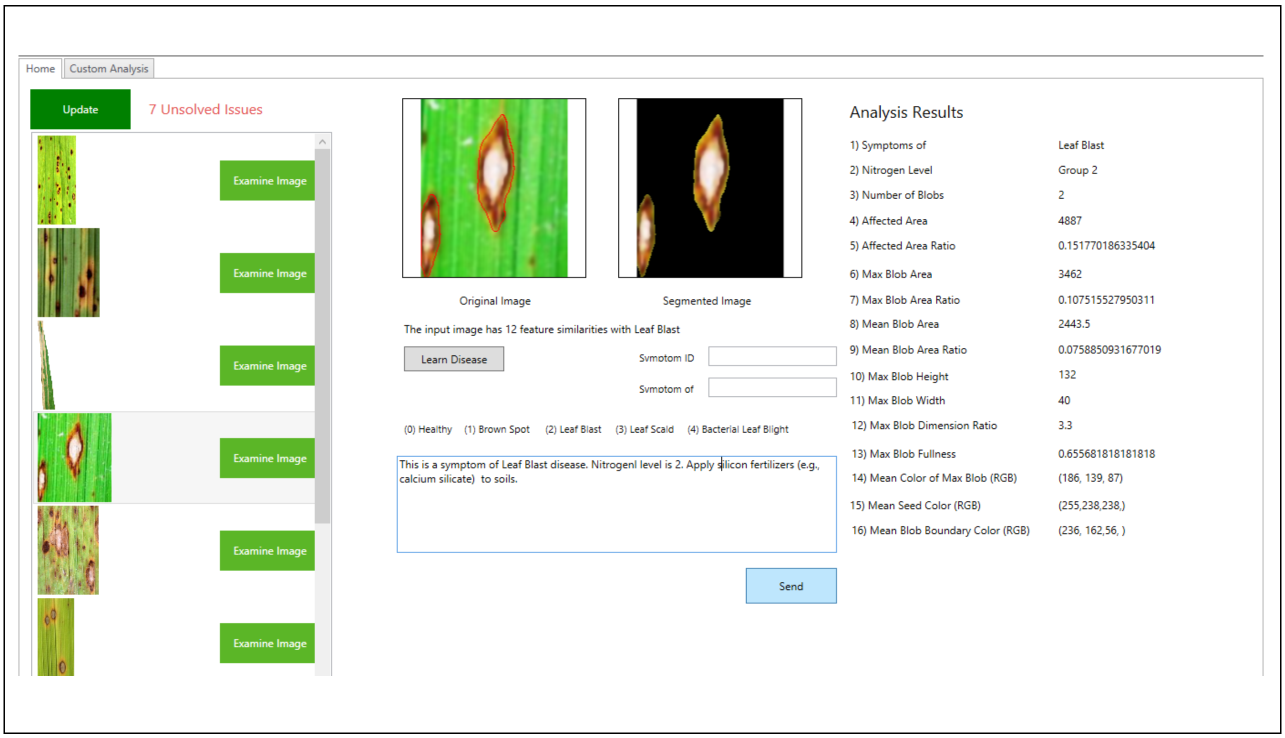


Figure 4.13: Presentation of classification results

4.3.6 Sending feedbacks

Every image in the database stores a device URL against it which is used for sending notification to the corresponding the farmer who posted the image for suggestions. A form is displayed in the software window where the expert can write his observations and suggestions and send to the farmer.

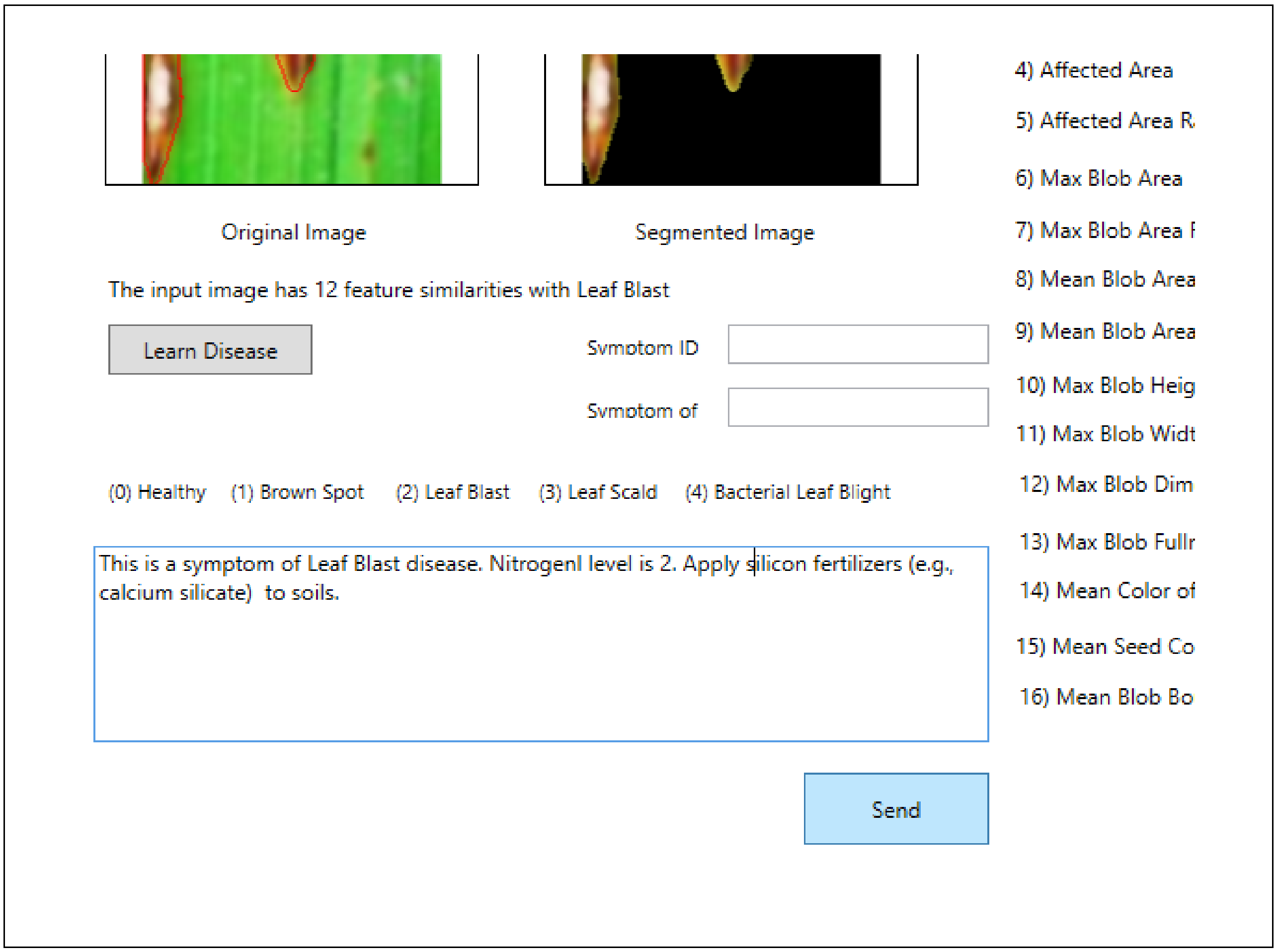


Figure 4.14: Sending feedback from the local server application

Once the farmer receives the notification, the receipt confirmation is displayed at the expert software panel.

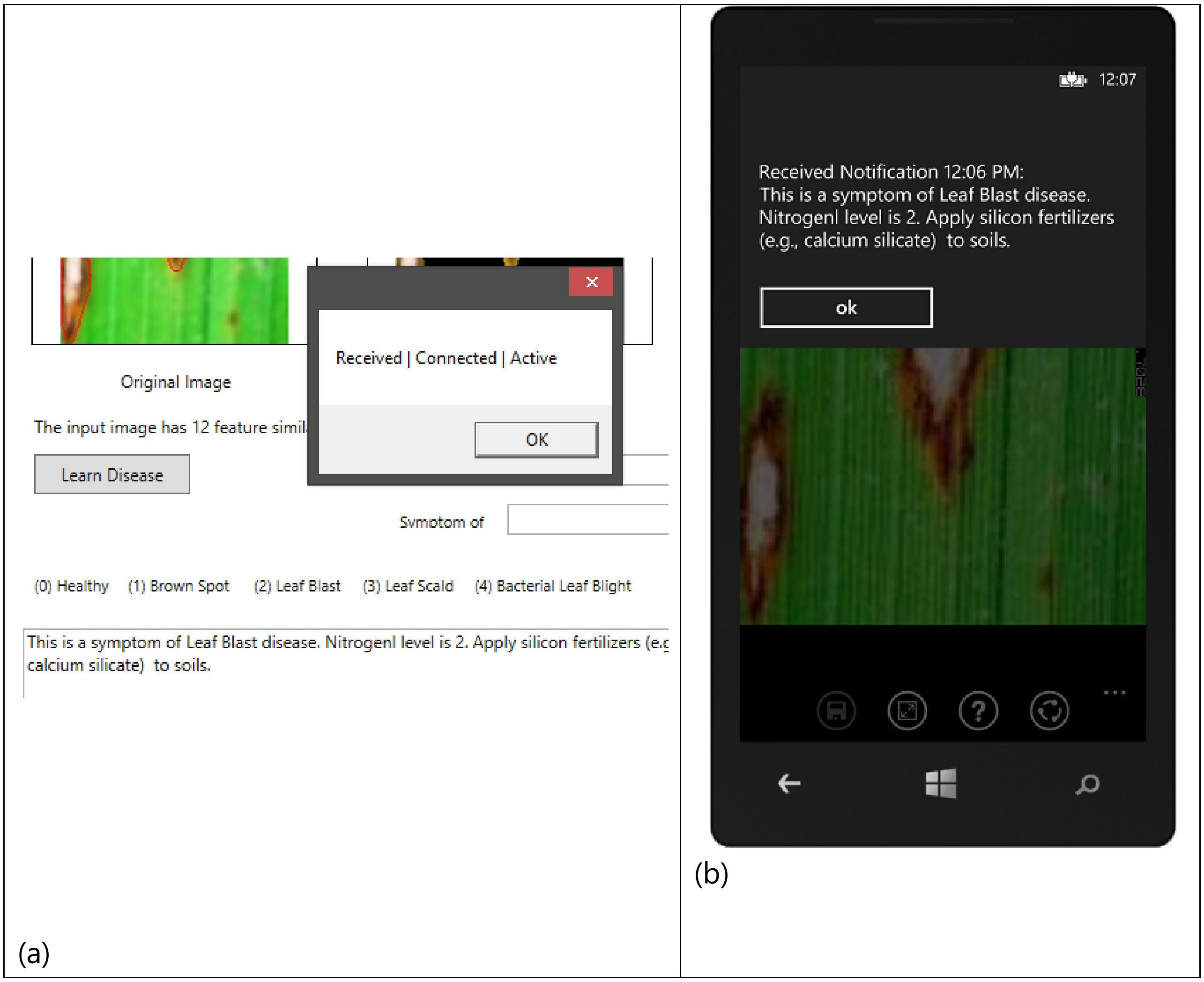


Figure 4.15: Notification receipt (b) and confirmation (a) view in expert panel

CHAPTER FIVE

EXPERIMENTAL RESULTS AND DISCUSSION

5.1 Results

We have considered very common paddy disease of the different areas over the world as experimental images. We have taken the images from the Lousiana State University Auricular Center (www.lsuagcenter.com), International Rice Research Institute (IRRI) and Bangladesh Rice Research Institute (BRRI). We have considered paddy disease images with the environmental parts. It has been observed that the proposed system output accuracy varies respect to paddy diseases.

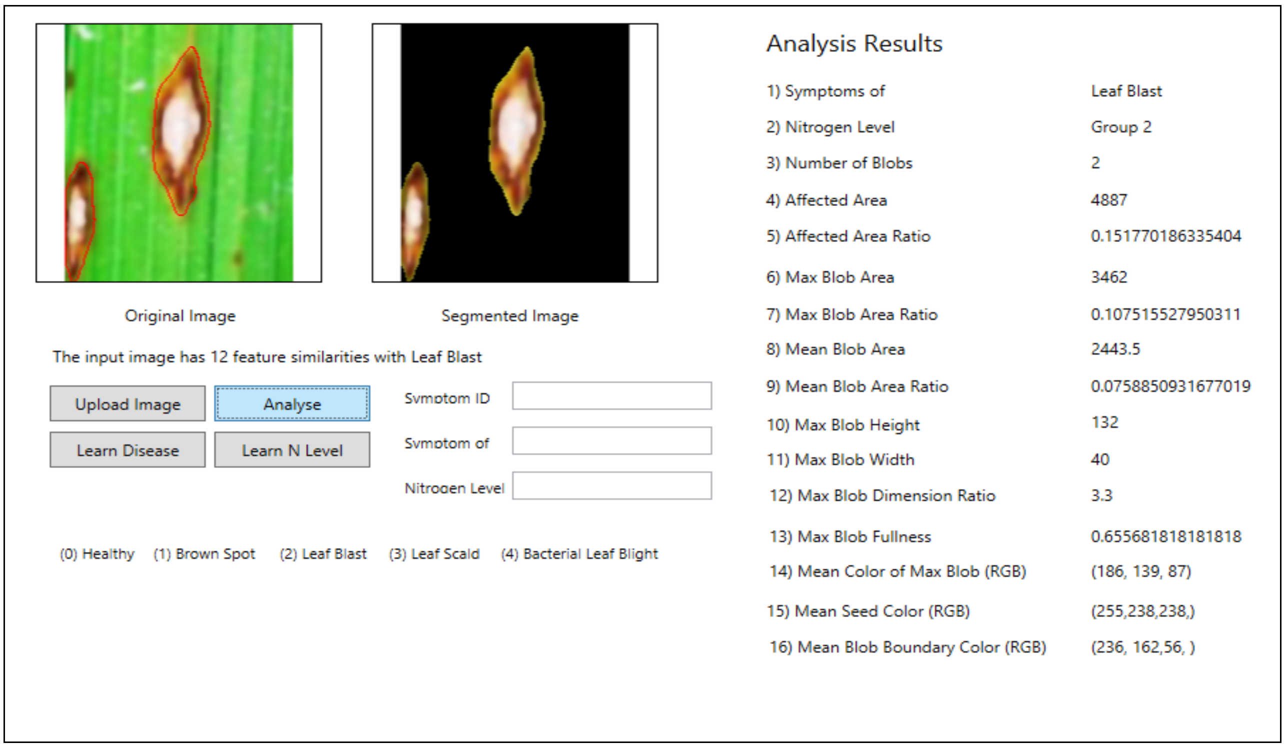


Figure 5.1: Paddy disease analysis results

To measure the performance of a paddy disease test, the concepts sensitivity and specificity are often used. Say we test some people for the presence of a disease. Say some of the test leaves have disease and we call it true recognition (*TR*) if the system recognizes the disease properly. In addition, if the system provides misleading results then it is called false recognition (*FR*). Thus, the number of true recognition and false recognition add up to 100% of the set. Finally we have calculated accuracy of the system using the following equation.

$$accuracy = \frac{TR}{TR + FR}$$

The following table depicts the data of accuracy for the targeted four paddy diseases analyzed by the system.

Disease Name	TR	FR	Accuracy (%)
Bacterial Leaf Blight	7	2	77.77
Brown Spot	29	3	90.62
Leaf Blast	17	1	94.44
Leaf Scald	11	1	91.66
Healthy paddy	23	23	100

Table 5.1: Comparative accuracy analysis of paddy disease recognition

Furthermore, training a disease using the system takes 575.0324 milliseconds. On the other hand, analysis of an image for disease detection takes 444.0253 milliseconds, which is an improvement from analysis time of 923.681 milliseconds for exhaustive comparison of features instead of using Binary Search Algorithm.

CHAPTER SIX

CONCLUSION AND FUTURE WORKS

5.1 Limitations

We have considered few limitations to the system. One of the drawbacks is regarding cancellation of background for image analysis of paddy disease. Analysis results of the diseases might not show accuracy in case of unwanted background of the picture. We dealt this situation by allowing the user to select only on the disease affected options using Nokia image zooming and crop capabilities. However, for diseases occurring in the tip or side of the paddy leaves, it would be necessary to consider applying background cancellations techniques In addition, the system is developed in English which might be challenging for the rural farmers to use. In order to actually spread this system for mass usage, it is essential that mobile application contains the instructions in Bangla.

5.2 Future works

We intend to continue updating this system for implementation of the project in real life. The primary focus will be to overcome the limitations of the currently developed system. Background cancellation algorithms need to be applied for the image processing to work for any image. In addition, we also want to upgrade this system that can detect any crop disease using only mobile phones.

5.3 Conclusion

Rice is an important food item not only for Bangladesh, but also for many other countries in the world. Alongside the supply of cultivation tools, the farmers also need access to accurate information that they can use for efficient crop management and there is no better way than providing them a service that they can use through their mobile phones. Although the system presented in this paper focuses only on four of the paddy diseases, 'Krishokbondhu' shows promises and gives us future direction for a robust application, making it a more effective tool that all farmers can use for management of all kinds of crops.

REFERENCES

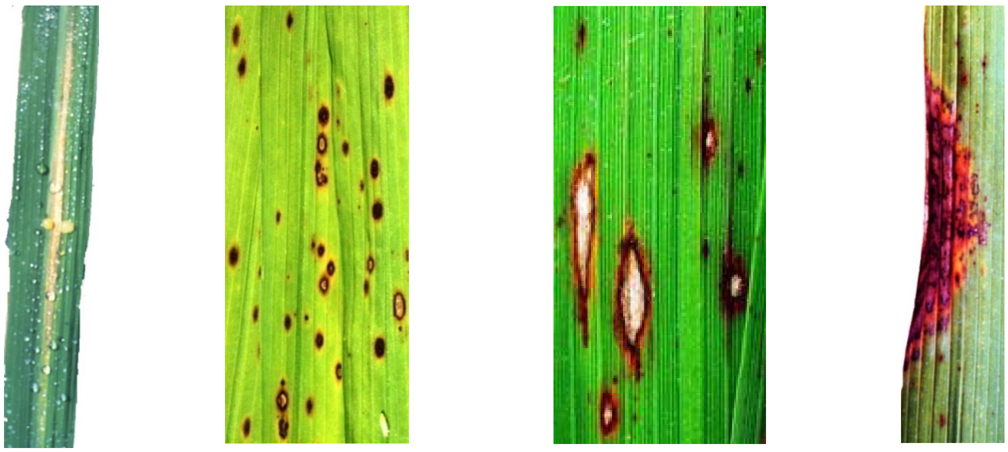
[1] http://www.knowledgebank.irri.org/step-by-step-production/growth/pests-and-diseases
[2] http://www.knowledgebank.irri.org/decision-tools/rice-doctor
[3] http://www.knowledgebank.irri.org/step-by-step-production/growth/pests-and-diseases/diseases
[4] http://en.wikipedia.org/wiki/Blob_detection
[5] http://msdn.microsoft.com/en-us/library/windows/apps/hh464924.aspx
[6] Sylvain Paris, Pierre Kornprobst, Jack Tumblin, Fred Durand "A Gentle Introduction to Bilateral Filtering and its Applications";
http://people.csail.mit.edu/sparis/bf_course/course_notes.pdf
[7] Lidi Wang, Tao Yang, Youwen Tian "Crop Disease Leaf Image Segmentation Method Based on Color Features", The International Federation for Information Processing, Volume 258, 2008, pp 713-717
[8] http://www.mathworks.com/matlabcentral/answers/110329-detecting-brown-spots-on-leafs
[9] Richard O. Duda, Peter E. Hart, David G. Stork. Pattern Classification, Second Edition
[10] D. Katsantonis, S.D. Koutroubas, D.A. Ntanos and E. Lupotto "EFFECT OF BLAST DISEASE ON NITROGEN ACCUMULATION AND REMOBILIZATION TO RICE GRAIN", Journal of Plant Pathology (2008), 90 (2), 263-272
[11] S. D. KoutroubasA, D. KatsantonisB, D. A. NtanosB,D and E. Lupotto "Nitrogen utilisation efficiency and grain yield components of rice varieties grown under blast disease stress", Australasian Plant Pathology, 2008, 37, 53–59
[12] Anne S. PrabhuI; Morel P. Barbosa Filho; Lawrence E. Datnoff; George H. Snyder; Rodrigo F. Berni; Fabricio A. Rodrigues; Leandro J. Dallagnol "Silicon reduces brown spot severity and grain discoloration on several rice genotypes", Tropical Plant Pathology, vol. 37(6):409-414, 2012
[13] http://www.knowledgebank.irri.org/step-by-step-production/growth/soil-fertility/leaf-color-chart
[14] http://www.aforgenet.com/framework/docs/html/c855bd08-b121-7233-cd64-294801f5a201.htm
[15] Shilpa Dantulwar, R.K.Krishna "PERFORMANCE ANALYSIS USING SINGLE SEEDED REGION GROWING ALGORITHM", International Journal of Innovative Research in Advanced Engineering, Volume 1 Issue 6 (July 2014), ISSN: 2349-2163

[16] Phadikar, S., Sil, J. "Rice disease identification using pattern recognition techniques", 11th International Conference on Computer and Information Technology (ICCIT),2008
[17] Reinald Adrian D. L. Pugoy, Vladimir Y. Mariano "Automated rice leaf disease detection using color image analysis", SPIE 8009, Third International Conference on Digital Image Processing (ICDIP), 2011
[18] Rahat Yasir, Nova Ahmed "Beetles: A Mobile Application to Detect Crop Disease for Farmers in Rural Area", Workshop on Human And Technology (WHAT), 8-10 March 2014, Khulna, Bangladesh
[19] P. Revathi, M. Hemalatha, Homogenous Segmentation based Edge Detection Techniques for Proficient Identification of the Cotton Leaf Spot Diseases. International Journal of Computer Applications (0975 – 888) Volume 47– No.2, June 2012 (2012)
[20] Amos Gichamba, Ismail A. Lukandu "A Model for designing M-Agriculture Applications for Dairy Farming", The African Journal of Information Systems, Volume 4, Issue 4, 2012

APPENDICES

A. Survey questionnaire on paddy disease visual symptoms (BRRI)

Target group of paddy diseases



(a) (b) (c) (d)

Figure: Variety in visual symptoms for different paddy diseases, (a) Bacterial leaf blight, (b) Brown spot, (c) Leaf blast, (d) Leaf scald

Method of analysis

Visual symptoms in the paddy leaf, especially features of the lesion.

Key features analyzed

- Nitrogen level
- Number of blobs/ spots
- Color for centre of the lesion
- Color for boundary of the lesion
- Disease affected area
- Ratio of affected region (with the entire area)
- Average blob area
- Biggest blob area
- Ratio of the biggest blob area
- Biggest blob height
- Biggest blob width
- Dimension ratio of the biggest blob
- Fullness of the biggest blob
- Average color value of the lesions

Questionnaire

- What is the possibility of detecting a paddy disease correctly observing only the visual symptoms of the leaf?

- Will the considered visual features work for all stages of the diseases?

- Does nitrogen level have any direct link with any diseases to add it as a feature for disease recognition?

- What other visual features might be considerable for development of such a system?

- Recommendations:

B. Feature values different paddy diseases

Disease Name	N Level	Spots	Area	Area Ratio	Max Area	Max Area Ratio	Height	Width	Dimension Ratio	Fullness	Mean R	Mean G	Mean B
Brown Spot	2	3	2601	0.162563	1681	0.105063	70	38	1.842105	0.631955	175	150	52
Brown Spot	3	8	2915	0.211232	1417	0.102681	81	29	2.793103	0.603235	130	98	42
Brown Spot	4	4	10828	0.436613	7129	0.28746	200	61	3.278689	0.584344	114	80	33
Brown Spot	3	12	4124	0.212577	2691	0.138711	126	43	2.930233	0.496678	169	140	62
Brown Spot	2	18	12629	0.474774	6433	0.241842	173	64	2.703125	0.581015	182	145	92
Brown Spot	2	22	8407	0.356229	2747	0.116398	135	35	3.857143	0.581376	173	137	87
Brown Spot	2	12	9031	0.586429	5655	0.367208	160	51	3.137255	0.693015	202	163	111
Brown Spot	2	9	8587	0.523598	3288	0.200488	88	55	1.6	0.679339	193	156	101
Brown Spot	3	6	2792	0.349	1214	0.15175	132	16	8.25	0.574811	151	113	54
Brown Spot	2	8	7854	0.478902	5104	0.31122	200	49	4.081633	0.520816	210	183	105
Brown Spot	2	8	1559	0.059504	384	0.014656	39	13	3	0.757396	146	128	16
Brown Spot	2	23	1102	0.066386	234	0.014096	42	8	5.25	0.696429	155	127	25
Brown Spot	3	7	7442	0.273603	4736	0.174118	137	55	2.490909	0.628534	131	96	41
Brown Spot	2	3	10950	0.55303	6097	0.307929	200	52	3.846154	0.58625	213	192	104
Leaf Blast	2	1	2472	0.152593	2472	0.152593	177	28	6.321429	0.498789	197	166	119
Leaf Blast	3	1	2818	0.10594	2818	0.10594	172	31	5.548387	0.528507	111	75	49
Leaf Blast	2	2	4887	0.15177	3462	0.107516	132	40	3.3	0.655682	186	139	87
Leaf Blast	2	6	2331	0.121406	894	0.046563	71	17	4.176471	0.740679	135	94	55
Leaf Blast	2	2	3897	0.192921	2053	0.101634	152	25	6.08	0.540263	188	151	102
Leaf Blast	2	4	7174	0.407614	5601	0.318239	200	55	3.636364	0.509182	166	140	63
Leaf Blast	4	2	4923	0.447545	3214	0.292182	138	33	4.181818	0.705753	124	90	45
Leaf Blast	4	1	3415	0.276787	3415	0.276787	167	33	5.060606	0.61967	148	119	73
Leaf Blast	4	3	7501	0.390677	7494	0.390313	200	78	2.564103	0.480385	164	136	95
Leaf Blast	4	5	3678	0.155847	3510	0.148729	200	32	6.25	0.548438	109	78	30
Leaf Blast	2	1	4396	0.180164	4396	0.180164	172	48	3.583333	0.532461	210	177	112
Leaf Blast	3	5	2115	0.203365	1108	0.106538	116	14	8.285714	0.682266	163	140	91
Leaf Blast	2	1	1960	0.110112	1960	0.110112	119	27	4.407407	0.610022	183	140	104
Leaf Blast	3	4	2279	0.151933	1824	0.1216	134	31	4.322581	0.439095	163	145	123

Leaf Blast	2	1	3682	0.239091	3682	0.239091	200	32	6.25	0.575313	148	129	79
Leaf Scald	2	4	4471	0.343923	4423	0.340231	163	49	3.326531	0.553775	158	117	52
Leaf Scald	2	2	5378	0.726757	5377	0.726622	187	37	5.054054	0.777135	170	127	98
Leaf Scald	2	7	4015	0.427128	3879	0.41266	180	41	4.390244	0.52561	157	92	85
Leaf Scald	3	1	4225	0.541667	4225	0.541667	190	34	5.588235	0.654025	162	141	89
Leaf Scald	2	2	5755	0.653977	5620	0.638636	200	37	5.405405	0.759459	125	73	45
Bacterial Leaf Blight	3	2	239	0.028452	219	0.026071	72	9	8	0.337963	221	214	159
Bacterial Leaf Blight	2	1	1792	0.242162	1792	0.242162	200	18	11.11111	0.497778	218	208	171
Bacterial Leaf Blight	3	3	854	0.129394	730	0.110606	130	8	16.25	0.701923	215	200	134
Bacterial Leaf Blight	3	4	2227	0.3275	2202	0.323824	200	22	9.090909	0.500455	206	181	153
Bacterial Leaf Blight	2	3	1208	0.1208	1203	0.1203	133	23	5.782609	0.393266	227	214	190
Brown Spot	2	3	10950	0.55303	6097	0.307929	200	52	3.846154	0.58625	213	192	104
Brown Spot	2	4	4782	0.11955	1706	0.04265	70	30	2.333333	0.812381	120	86	20
Bacterial Leaf Blight	3	9	4175	0.549342	3732	0.491053	200	26	7.692308	0.717692	205	166	112
Bacterial Leaf Blight	3	1	4254	0.506429	4254	0.506429	200	42	4.761905	0.506429	193	158	105

C. Nitrogen level color information in different intensities

N Level	Mean L	Mean A	Mean B
2	147.4535	108.0452	167.2717
2	161.4176	105.9169	179.3068
2	145.0581	105.2921	174.4218
2	183.1542	111.9156	165.1994
2	152.5852	102.9478	179.2504
2	139.2439	105.8263	167.048
2	144.9396	102.32	178.132
2	140.0224	102.6305	177.7698
2	126.0504	105.963	177.1827
3	136.9747	110.1459	157.6732
3	142.1977	105.6119	171.0513
3	130.9089	107.652	163.9232
3	168.0474	112.4372	157.4569
3	138.2968	102.8282	171.1144
3	123.526	109.5156	155.1098
3	130.6328	105.8313	168.0787
3	131.3594	104.14	169.562
3	120.3719	99.84365	169.5241
4	114.2989	116.3568	141.426
4	125.3485	111.3208	151.7278
4	100.9748	112.569	152.5193
4	142.449	115.8086	144.3967
4	120.6667	110.6966	151.2952
4	98.08005	116.9223	141.8518
4	105.8536	111.1444	153.3581
4	114.2619	107.8957	159.099
4	96.17228	102.547	159.7107
5	99.16563	120.9901	134.7529
5	89.76543	116.9763	139.6261
5	84.95665	117.3638	140.6828
5	117.9106	120.6377	134.571
5	79.5993	114.6714	140.7797
5	59.17105	120.8829	135.6549
5	82.60733	117.2428	143.4219
5	92.26773	112.8587	147.1167
5	83.88748	104.1416	147.1294