Early Detection and Classification of Paddy Diseases with Neural Networks and Fuzzy Logic

MOHD ADZHAR ABDUL KAHAR, SOFIANITA MUTALIB, SHUZLINA ABDUL-RAHMAN
Faculty of Computer and Mathematical Sciences
Universiti Teknologi MARA
40450 Shah Alam
MALAYSIA
j.adzhar@gmail.com, {sofi, shuzlina}@tmsk.uitm.edu.my
http://fskm.uitm.edu.my

Abstract: - This paper presents an integrated method for recognizing diseases on paddy plant leaves and provides the user with recommendations on how to overcome and control the diseases. The study focuses on three main paddy diseases in Malaysia that commonly affect the yield of crops namely Bacterial Leaf Blight (BLB), Leaf Blast Disease (LBD) and Bacterial Sheath Blight (BSB). The method of recognition used in this study is a neuro-fuzzy expert system. The method combines the learning capabilities of artificial neural network with the human like knowledge representation and clarification abilities of fuzzy logic systems and also a rule-based expert system. In this study, the prototype is developed to assist Malaysian paddy farmers and paddy researchers by providing an early recognition of diseases in their paddy crops and assisting them with their crop management. Effective crop management is vital to ensure the crops health and also to get a large amount of yield. The accuracy result for the recognition is 74.21%. Hopefully the findings in this study may assist farmers and researchers in managing and recognizing the early stage of their paddy diseases.

Key-Words: - Neural Network, Expert System, Fuzzy Logic, Crop Management, Paddy Diseases

1 Introduction

Malaysia is an agriculture based country that has many people working in the agriculture industry [3]. The agricultural sector plays an important role in Malaysia's economic development by providing rural employment [4]. According to TradingEconomics.com [15], as of January 2012 there are approximately 12.6 thousands workers in the sector. The industry functions in producing many crops, while paddy is the third most widely planted crop in Malaysia after oil palm and rubber [1]. Paddy is one of the nation's most important product as it is considered to be one of Malaysia's staple food and cereal crops and because of that, there has been many efforts done to ensure its safety, one of them is crop management of Paddy plants.

Most Malaysian paddy farmers currently lack the knowledge of detecting paddy diseases effectively. Presently, the conventional method hires experts to check on paddy fields, which consumes huge amount of time and money. Large numbers of experts are normally required to analyse whole paddy fields. Thus, to save time and money, experts normally generalize a particular area that may be infected. Moreover, delay in disease detection implicates some plants fail to produce yield which otherwise can easily be treated if the disease are detected earlier. It is hard for paddy farmers to recognize paddy diseases as most farmers lack the expertise know-how knowledge on paddy diseases. In addition, all paddy plants may not have same diseases. If the affected paddy plant is not cured rapidly, decrease of paddy production is the crucial concerns. Therefore this study aims to solve these problems by developing an integrated method in recognizing paddy diseases at their early stages and give recommendations on how to solve and control it.

2 Related Works

Crop Management can be defined as the method of agricultural practices used to enhance the development, growth and yield of crops. There are a few ways to manage paddy disease. The best solution is to prevent the disease from actually happening but in the event that it does happen to the crops, the Malaysian Agricultural Research and Development Institute (MARDI) would be providing solutions and recommendations for paddy farmers in Malaysia. The International Rice Research Institute (IRRI) has also provided a tool for crop management which they call “Crop Manager”and it is available at their website but unfortunately crop management for Malaysian paddy plants are not available currently. There are three main diseases that affect paddy plants namely Bacterial Leaf Blight (BLB), Leaf Blast Disease (LBD) and Bacterial Sheath Blight (BSB). Crop management is one of the most important aspects in farming. It refers to the different processes connected towards the effective development and harvesting of crops. It is vital to ensure the plants health in order to produce yield. Rate of crop illness and irritation is one of the real concerns faced by farmers [8]. Therefore maintaining healthy crops that are disease free is their highest objective. According to the International Rice Research Institute [6], there are over 99 paddy defects in different ranges from paddy diseases, pests and many more. This study will be focusing on three paddy diseases as explained in the next subsections. Over the years, the disease became more widespread and treated as global threat because of wide scale planting of a susceptible rice variety which then caused an estimated loss of approximately RM 50 million throughout 1982 to 1994 [13].

2.1 Bacterial Leaf Blight

Bacterial Leaf Blight (BLB) is caused by *Xanthomonas oryzae* pv. Oryzae [6] causes wilting of the seedlings and yellowing and drying of the leaves [6]. BLB was initially located in the paddy fields of Peninsular Malaysia during the early 80s, however it was in a limited widespread [13].

BLB is considered to be one of the most serious diseases of paddy plants. The earlier the disease happens, the higher the yield loss. Yield loss due to this disease might be as high as 70% when susceptible varieties of rice are grown, in situations favourable to the disease [14]. There are 3 main stages for BLB and it can be referred to in the Standard Evaluation System for Rice Book [5]. According to the SES [5], early stages for BLB are from 1 to 4, middle stages are from 4 – 6 and late stages are from 7 – 9. The symptoms of BLB are when there are yellowing on the leaves. The disease affects in the middle stages when wilting of the leaves have occurred and when the wilting has covered 50% or more of the leaf area, it has entered into the late stages of the disease.

There have not been BLB cases in Malaysia for almost 10 years, but it has re-emerged recently. In serious cases, BLB can affect as much as 50% to 70% of the amount of yield product [7] therefore it is important for the paddy farmers and researchers of BLB to know the solution to this problem. One of the ways to manage BLB is by planting resistant varieties [7]. This method has been proven to best, reliable and cheapest way to control BLB [6]. Another important controlling method is to destroy the source of the disease or pathogen. Other option includes using a balanced number of nutrients for the plants such as nitrogen and also an optimum amount of seeds which is approximately 120 kg / hectare [7] and to change the water in the plant field. Using an unbalanced amount of nitrogen can cause the plant to have weak silica and using too much seeds will expose the plant to other health problems. BLB can also be controlled by the usage of pesticides. Pesticides that are specially designed for bacterial plan infections are copped based and it is very toxic [7]. These pesticides cannot be combined with other types of pesticides as it could be more hazardous to the plant than it already is. With the application of pesticides, the amount and timing factor of using the pesticides is crucial. Using too much pesticide can result to the produce being dangerous for consumption.

2.2 Leaf Blast Disease

Leaf Blast Disease (LBD) is caused by the fungus *Magnaporthe oryzae* [6]. LBD is a primary issue in the rice development process [9] as it could occur all year long and also to every country that plants paddy crops [7]. There are three ways the disease could be spread which are through air, seeds and also alternative plant host [7]. LBD can affect all above ground parts of a rice plant which includes the leaf, neck, collar, node, parts of panicle, and sometimes leaf sheath [6]. LBD usually occurs after 15 days of planting the seeds [7] and also can be classified into three different stages, which are early, middle and late stages. The stages could be identified by referring to the SES [5] or to an expert on the disease.

Leaf Blast Disease (LBD) is also considered to be one of the most dangerous and popular paddy disease to occur in Malaysia [7] as it usually attacks the plant as early as 15 days after planting and it could affect as much a 100% of yield loss. LBD is taken very seriously and can only be controlled during its early and middle stage [7]. Once LBD has reached its middle stages, there is only a 50% chance of the plant surviving and when it has gone into its late stage, there is little to no chance of the plant to survive. The primary control option for LBD is by planting resistant varieties [6]. Other management options can also be done such as adjusting the planting time, using optimum number of seeds (120 kg / hectare), and also the right amount of plant nutrients such as nitrogen. Applying too much nitrogen can cause the plant to develop weak silica. LBD can also be controlled by the usage of pesticides. Pesticides such as tricyclazole, isoprothiolane and azoxystrobin can be used to control LBD. For Malaysian paddy farmers, these pesticides are provided by the government as subsidies and they also provide guides on the usage of the pesticides. Once LBD is detected, it is suggested to spray the pesticides for 3 days or spraying the pesticides before the collar of the plant blooms (within 70 days after planting). According to SES [5], early stages for LBD are from 1 to 4, middle stages are from 4 – 6 and late stages are from 7 – 9. The symptoms of LBD are lesions on the leaf. It usually begins with white to gray–green spore lesions or spots that resembles a diamond or bird eye shape [7] with dark green borders. When the borders of the lesions have started to dry or changed color to a yellow brown color, it signals that the disease has moved on to the middle stage of the disease. The late stage of LBD can be seen when the lesions have covered 50% and more of the leaf area.

2.3 Bacterial Sheath Blight

Bacterial Sheath Blight (BSB) is caused by *Rhizoctonia Solani* [6], a type of fungal disease that causes the infected leaves to rapidly dry out. According to Jack [7], BSB is similar to BLB as it is both a bacterial type of disease. Paddy plants can be infected with BSB by a number of ways such as through water used, the ground, infected from other plants and also alternative hosts but unlike BLB and LBD, BSB cannot be infected through air-born spores [7]. According to Jack [7], BSB usually occurs at either the leaf or the stems of the paddy plant and they usually occur from the 30th – 40th – 55th day after planting the seeds of the plant. Early stages for BSB are from 1 to 4, middle stages are from 4 to 6 and late stages are from 7 to 9 which can be referred to in SES [5]. One of the symptoms of BSB is that the disease has an ellipsoidal or an oval shaped greening to gray lesions on the plant [6]. It usually begins with a small white to gray–green lesions that resembles oval or ellipsoidal shape [7] with dark green borders. When the lesions have started to expand and combine with each other, it signals that the disease has moved on to the middle stage of the disease. The late stage of BSB can be seen when the lesions have covered 50% and more of the leaf area.

Bacterial Sheath Blight (BSB) is one of three of the most popular and dangerous paddy diseases to occur in Malaysia [7]. BSB is considered dangerous because it can affect a wide area of the crop once an area of the crop is infected. With BSB, the infection usually begins at the roots and stems of the plant. When the leaf of the plant is started to have lesions, it is very important to take action before the plant becomes damaged. According to IRRI [6], currently there are no paddy varieties that are resistant to BSB available for cultivation. The first option in controlling BSB is to destroy the pathogen or source of the infection. Once the source is destroyed, it will be easier to control it. Other options include using a balance amount of seeds and also nitrogen, control of weeds (mainly focus on the leaves), drain the crop fields in the early stage of the cropping season to reduce the amount of BSB epidemics [6]. Pesticides can also be used to control BSB such as *difecnoconazole*. This pesticide is effective after the disease has been identified and confirmed and the right amount of usage is also important.

2.4 Diagnosing Paddy Disease

A research done by Kurniawati [9] to develop a diagnosis system mainly based on MATLAB to recognize paddy diseases which focuses on Blast Disease (BD), Brown-Spot Disease (BSD) and Narrow Brown-Spot Disease (NBSD). The research concentrates on extracting paddy features through off-line image which involves converting RGB images into a binary image using variable, global and automatic threshold based on Otsu method. The authors used a morphological algorithm to remove the noises in the images by using region filling technique. Then the image characteristics of the paddy leaves consisting of lesion percentage, lesion type, boundary color, spot color and broken paddy leaf color are extracted. Kurniawati [9] stated between the three thresholding methods that have been applied, variable threshold method has the highest accuracy results at about 87.5%. The author suggested future works to focus on ANN to classify paddy diseases. Anthonys [2] stated that the classification and recognition of paddy disease are a major technical and economical importance in the agricultural industry. The goal of the research was to develop an image recognition system that is able of detecting paddy diseases. The research focuses on three major diseases commonly found in Sri Lanka, which are Rice Blast, Rice Sheath Blight and Brown Spots. In their research, the images were acquired under a controlled laboratory condition using a digital camera with 50 sample images. Anthonys [2] starts the image processing phase by first digitizing the color image of the paddy leaf. Figure 1 shows the original image and the digitized image of the paddy leaves. Anthonys [2] stated that an image image captured by a sensor is expressed as a continuous function *f(x,y)* of two co-ordinates in the plane. The authors also explained that image digitization means the function *f(x,y)* is sampled into a matrix with M rows and N columns.

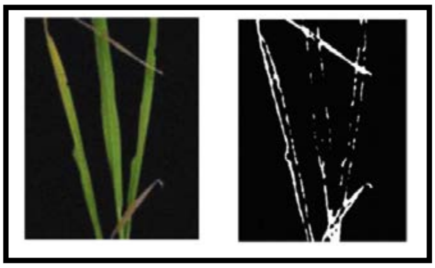


Figure 1 Image digitization results [2]

The next step conducted in the research was to segment the images which are also the first step in analyzing the images. The research used Soble edge detection to detect the edges of the images. The color textures of the paddy leaf in the images are also extracted. Results of the segmentation are shown in figure 2 and the results of the extraction of color texture are shown in figure 3.

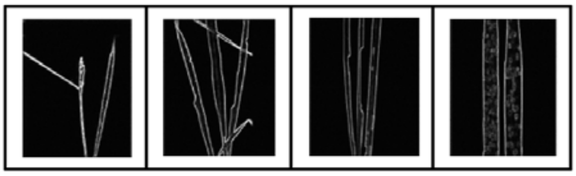


Figure 2 Image segmentation results [2]

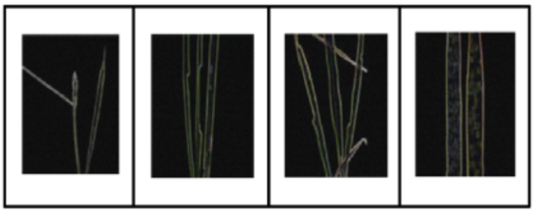


Figure 3 Extraction of color texture results [2]

The authors extracted the color features of the paddy leaves by using CIE L*a*b* color space where the authors stated the L*a*b* color consists of luminosity 'L' or brightness layer, chromaticity layer 'a' and 'b'. The research used Euclidean distance to measure the difference of the two kinds of color. The color values are then displayed on a scatter plot to show the difference between the colors. Figure 4 shows a sample of a scatterplot of the segmented pixels in 'a*b*' space.

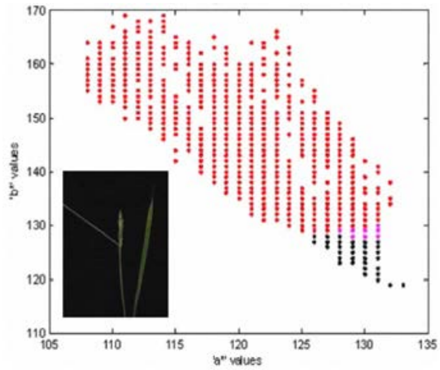


Figure 2 Scatterplot of the segmented pixels in 'a*b' space [2]

The research then used nearest neighbour classification to classify the membership function for each class. The format for the membership function in the research is *MF (RGB range, shape, length, width, diameter . . .)*. The final step was the recognition phase of the diseases. The system also provides the user with the signs and symptoms for each type of disease to give a better accurate diagnosis of the disease. The results of the system was 80% for Rice Blast (20 images), 60% for Rice sheath blight (15 images) and 85% for Brown spots (11 images). The authors also suggested in using a camera with a stronger color sensor to allow higher precision of the image color and segmentation.

Another study conducted by Nooro [11] that focuses on recognizing different type of herb shrubs that are available in the forest by using image recognition of shrubs leaves and Back-propagation Neural Network. According to Nooro [11] by being able to recognize the different types of herbal shrubs, it could benefit hikers as it could help them in case of emergency. In this study, the author uses the leaf feature of the herbal shrubs to recognize the type of shrubs. The author extracts the features of leaf using MATLAB. Table 1 and table 2 shows the steps taken by the author to extract the features of leaf using MATLAB. After the features are extracted, they are converted into binary data and it will be save in a .CSV data file which will be later be fed to the Back-propagation Neural Network Engine of the system.

The highest accuracy for the system was 80% which was higher than the expected result from the author which was 70%. Nooro [11] suggested a combination of other techniques with Back-propagation Neural Network could be used to increase the accuracy rate of the recognition of the type of shrubs. Also the images should be acquired

ISBN: 978-1-61804-302-3

in a controlled environment to smoothen the image pre-processing phase.

Table 1 Steps taken to pre-process the leaf shrub image [11]

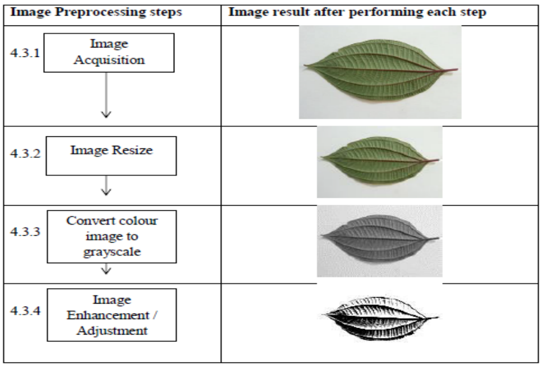
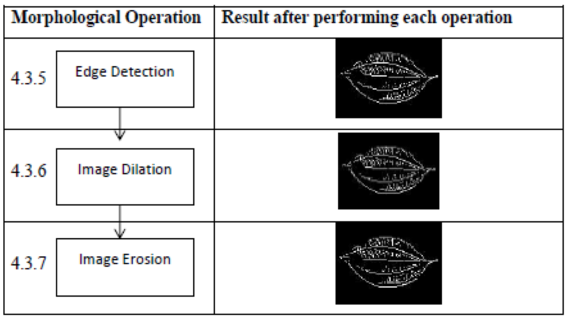


Image Preprocessing steps	Image result after performing each step
4.3.1 Image Acquisition	
4.3.2 Image Resize	
4.3.3 Convert colour image to grayscale	
4.3.4 Image Enhancement / Adjustment	

Table 2 Steps taken to extract the features of the leaf shrubs [11]



Morphological Operation	Result after performing each operation
4.3.5 Edge Detection	
4.3.6 Image Dilation	
4.3.7 Image Erosion	

3 Methods

The images were collected from MARDI research station in Bertam, Pulau Piang, IRRI and the features of disease data were collected from MARDI, IRRI, and other resources. The images that were collected from MARDI are used in this work while the other images are only used for references only. The images were collected using a Canon 550D with 2 lenses, Canon 50mm 1.8 F and also Canon 28 – 80mm 4 – 5.6 F lenses. During the time of the attachment, there were four available paddy diseases to be collected. The images were acquired in the paddy field while the leaf was still intact with the plant. It was done to replicate paddy plant with disease in real world conditions. The total numbers of the acquired images are stated in Data Set Management. Table 3 shows the number of images that have been collected. The images are organized in terms of its type and also the stage of the disease. The unusable images are those that are either blurred images or images that have bad lighting that cannot be used in the project.

Table 3 Total image data collected that have been filtered and evaluated

Defect Type / Stage	BLB	LBD	BSB	Brown Spots	Healthy Plants
Early	281	52	22	118	-
Middle	71	164	51	54	-
Late	22	110	31	61	-
Unusable	35	34	15	86	-
Total	**409**	**360**	**119**	**319**	**239**

The collected data undergo a pre-processing stage that uses MATLAB to read the RGB image. The next step is to crop the image. Here, the user can select which part of the image to be recognized in the system. Since the images collected were not uniformed, the images have to be cropped manually. The third step is to resize the cropped image. Here, the image is resized to a dimension of [50 X 50]. This is done in order to make the data uniform to be fed into the system.

3.1 Feature Extraction

After the images are pre-processed, the features of the paddy disease need to be extracted to be fed to the system for recognition which is the shape of the disease. This process is also done in MATLAB which uses morphological operations. The feature extraction of the image is done from the continuous steps of pre-processing the images. First step is the process of converting the images from RGB color profile to Grayscale color profile. Once done, the next step is to detect the edges of the defect. Canny Edge Detection was used in this project. A series of experiments were conducted that lead to decision of using Canny Edge Detection in this project. The last step in this stage is smoothing the image. This is done to enhance the image's features. In this step, a combination of image dilation and erosion was used.

3.2 Color Processing

The last stage in the pre-processing of the image is the processing of the color from the defects in the images. This step can be done both before and after feature extraction. The colors of the defects are processed in order to determine the stage of the disease. The scope of the project contains 2 stages, which are the early and middle stages of the disease. The color of each stage of the disease are different, therefore by processing the color of the disease it is possible to differentiate the stages. The color processing begins with converting the resized image from the pre-processing stage from RGB color profile to a LAB color profile image. Then, to display the data of the LAB color, a scatter plot was

ISBN: 978-1-61804-302-3

used. This is done by referring previous works and trying to replicate the project.

The first step is converting the RGB color profile to a LAB color profile. The second step is the color descriptor. The term color descriptor is used when displaying the color of an image in a readable data form [12]. To descript the color of the disease in this project, it is broken down into a few steps, first is to retrieve the data from the A* and B* values from the LAB image, then is to remove the zero values, and lastly is to display the data on a scatter plot. The system will start with an image of a paddy plant, then the image will be pre-processed using MATLAB. The features extracted will consist of color and shape of the paddy plant with disease. Next is the training process and the testing process for the systems accuracy. After that is the recognition of the plant disease and lastly is classification of the plant defect.

3.3 Architecture Design

Figure 5 displays the systems architecture design. The system uses a neuro fuzzy architecture which consists of a back-propagation neural network combined with fuzzy logic. The images of paddy leaves are first acquired and then pre-processed in MATLAB. Then features such as color and shape of the leaf are then extracted. The features are then inserted into the systems engine, color of leaves are inserted into the fuzzy logic engine and the shape of leaves are inserted into the BPNN engine. The results are then evaluated to find the best weighted value to be used for system testing. After the recognition part is done, the results will be matched in the database to classify the paddy disease with its solutions.

3.4 Back-Propagation Neural Network Architecture Design

In this study, the Back-Propagation Neural Network (BPNN) part of the system is used to determine the type of disease with the process of training and testing to reduce the amount of errors in the network. After the training is completed, the network is tested using the testing data to produce an output of the type of disease. The BPNN architecture is divided into three layers, which are the input layer, hidden layer and also output layer. 2500 input nodes are used in the input layer. The 2500 nodes came from the dimension of the processed image which is 50 X 50 pixels of data. The hidden node used is 1250 nodes and the number of nodes in the output layer is three nodes. In this architecture, there are three output nodes represent paddy diseases to be recognized which are BLB, LBD and BSB. Figure 6 illustrates the BPNN architecture for this project.

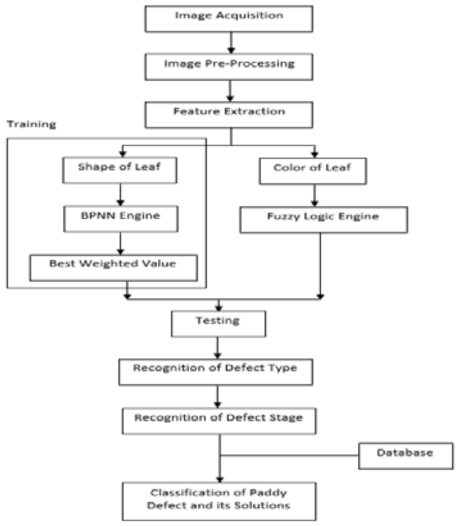


Figure 5 Architecture design

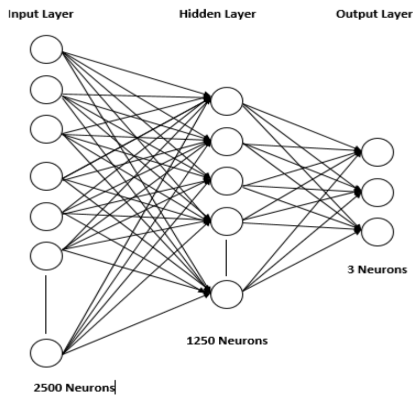


Figure 6 BPNN Architecture

The training part of the network involves four stages which consist of feed forward calculation, back-propagation of the error and lastly the adjustment of weights. In the feed forward phase, the input layer receives input signal before being transmit to the hidden layer and then computes its activation. The signal was sent to the output layer before the activation was computed. The output layer also made a comparison between the activation that was computed and the target data set before determining the amount of error. The error is then distributed back to the previous layers for the adjustment of the weights and this process was repeated until the stopping conditions were met.

ISBN: 978-1-61804-302-3

3.5 Fuzzy Color Processing

The color features in the image were processed as discussed in section 3.2. After the color of the disease is pre-processed, the images are represented in the form of decimal numbers for both values of A* and B*. By using the function "*mean()*" in MATLAB, the system was able to retrieve the mean or average of the A* and B* values. The A* and B* values were firstly analyzed and only non-zero values which was ranging from 1 – 256 were used in descripting the color on the scatter plot. Then, only the non-zero values were used in calculating the mean of both A* and B* values. MATLAB function "*mean()*" was used where it returns the mean value of the variable it is intended for, in this project, the A* and B* values. The mean values are then saved into separate .CSV file to be used in the system later for recognizing the paddy disease's stage. The data files are then classified into its disease types and stages. The stages of the diseases can be classified by using the formula below.

$$Stage\ Membership\ Class = \frac{Total\ added\ mean\ values}{Number\ of\ mean\ files\ used} \quad (1)$$

By using the equation above, the stages for the disease can be classified by calculating it with the mean of A* and B* values from before. For an example, if there are 10 mean data files for values of A* and B* to be used for a type of defect at a certain stage, then all 10 mean data files will be added up and divided by the number of files used. After calculating for each stage class for each type of the disease, the membership classes were acquired and recorded in Table 4.

Table 4 Membership class for paddy disease stages

Disease Type	Membership Class			
	A* Values		B* Values	
	Early	Middle	Early	Middle
Bacterial Leaf Blight (BLB)	0 – 126.00	> 126.00	0 – 167.00	> 167.00
Leaf Blast Disease (LBD)	0 – 112.00	> 112.00	0 – 157.00	> 157.00
Bacterial Sheath Blight (BSB)	0 – 162.00	> 162.00	0 – 120.00	> 120.00

4 Prototype Development

The development process is divided into two stages. First stage is to program the backend of the program. For the ANN development phase, JAVA is used and for the pre-processing phase will be using MATLAB. The second stage is to conduct the frontend of the prototype using NetBeans for making the interface of the system and the database and MATLAB GUI BUILDER to make the interface for the image pre-processing GUI.

4.1 Preprocessing Architecture

The system uses images to recognize paddy diseases. All of the raw images collected have been pre-processed in order to be fed to the system to be processed. Figure 6 shows the flowchart for the image pre-processing process. It first starts off with a few basic steps then it is divided into 2 parts which is for the BPNN engine and also the Fuzzy Logic Engine. The next image segmentation experiment conducted used the Region of Interest (ROI) Poly feature. The ROI Poly enables the user the select which part or region of the image to be segmented and the rest of the image will be removed. This experiment seemed to be the most suitable for this research as it could segment the region of interest from the background cleanier.

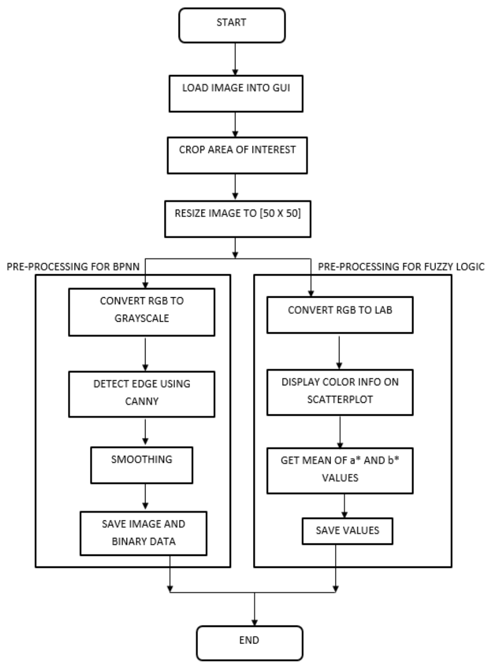


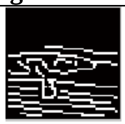
Figure 6 Image Preprocessing Architecture

ISBN: 978-1-61804-302-3

4.2 BPNN Image Data

There were a few experiments conducted for the pre-processing of the BPNN image data. The first experiment conducted was to compare the difference in terms of edge detection and also image enhancement. Table 5 shows the comparison between edge detection techniques. In this experiment, a comparison between edge detection techniques was done to determine which technique will be most suitable for this project. It was concluded that Canny Edge Detection was the most suitable due to the fact it could detect in more detail of the defect features on the plant leaf.

Table 5 Edge Detection Comparison

Edge Detection Comparison			
			
Roberts	Canny	Sobel	Prewitt

4.3 Color Processing Image Data

The data to be pre-processed for the engine that differentiates the stages are from the color features of the leaf. The color features are extracted through plotting a scatter plot. After the image containing the region of interest has been resized, the image is then converted from RGB color profile to LAB color profile where L is the lightness or brightness of the image, A and B are the color-opponent dimensions based on nonlinear compressed coordinates. An experiment was conducted in this section to determine the size of the image to be used in the system. There is a big difference between the 2 scatterplots, where the image with the size [50 X 50] shows a measurable scatterplot that can be clearly seen and the image with the size [30 X 30] has a scatter plot that is hardly able to be seen and also measured. With this experiment, it was concluded that the minimum size for the image to be used is [50 X 50].

4.4 Experimental Setup

The training process is crucial for the system in order for it to be able to recognize paddy disease and be able to deliver the desired output. In this project, the system was trained to recognize the type of paddy disease whether it is BLB, BSB or LBD. For BLB there are 28 images for BLB at its early stages and there are 28 images at its middle stages. For LBD there are 33 images in its early stages, and 43 images in its middle stages. Lastly, there are 32 images for early stages of BSB. The total number of images is 164. The ratio of data used in the system is 80:20 where the data for training will be 80% which is 133 and data for testing will be 20% which is 31 images. Table 6 shows the initial parameters used for training the system.

Table 6 Initial parameters used in system training

Parameter	Description
Input values	Binary
Number of input neurons	2500
Number of hidden neurons	500
Number of output neurons	3
Learning Rate	0.2
Activation Function	Sigmoid
Stopping Condition	Epoch < 10000 & MSE < 0.0001

The training process begins with receiving binary image data which is in an excel file format (.CSV). Then it calculates the sum of the weight input by each hidden neuron and sends it to the output layer. Each sum of the weighted input will be applied an activation function to compute the output signal. Next, the error was computed by differentiating between the desired and actual output of the output neuron. The error is then sent to the layer below associated with new updated weight before each hidden neuron totals up its delta weight and calculates its weight correction for the input neuron and the weight will be updated for each output and hidden neuron, then the stopping condition is tested. The network learns by the error for every epoch for this leaning process. The output from the network was used for recognizing the type of paddy defects.

The system is trained using different numbers of learning rates ranging from 0.1 to 1.0. This was done in order to identify the optimal learning rate needed for the training part of the system. The system was experimented using a series number of learning rates to test for optimal value. In this experiment, the number of hidden neurons used was 500 nodes and the number of maximum epochs used was 30000 as stated in the initial parameters. After the experiments were completed, it was found that the optimal learning rate to be used is 0.8 with an accuracy of 100%, number of epochs was 110 and MSE was 0.000003.

ISBN: 978-1-61804-302-3

5 Results and Discussion

5.1 Recognition of Disease Type

In this phase, the system will be tested for correct recognition of the paddy defect types. The testing data comes from a binary image data file (.CSV) that has been pre-processed earlier. The accuracy of the correct data can be calculated after the testing process is completed. The accuracy of the system is expected not to be 100% due to the performance of the recognition and the data used. The system is expected to get about 60% of accuracy for the system testing process. This is due to the fact that the images were not captured in a controlled environment and there are noises in the image that could not be removed. Therefore the images have noise that interrupts the data and would make the system difficult to recognize. The accuracy of the system is calculated using the formula in the equation below.

$$Accuracy\ Result = \frac{Number\ of\ correct\ samples}{Number\ of\ total\ samples} * 100\% \quad (2)$$

5.2 Recognition of Disease Stage

In this phase, the system will be tested for correct recognition of the paddy defect stages. The testing data comes from a binary data file (.CSV) that contains the mean of A* values and B* values that has been pre-processed earlier. The accuracy of the correct data can be calculated after the testing process is completed. The accuracy result will still use the same formula as in equation 2. After the system recognizes the correct paddy disease stages, it will be pass on the data to the next interface where all the data recognize will be shown and solutions for the disease is also shown.

Table 7 Accuracy Results for Recognition of Disease Stage

Type	**Number of Correct Recognition**	**Number of Incorrect Recognition**	**Accuracy (%)**
BLB	6	6	50.00
BSB	8	1	88.88
LBD	15	1	83.75
Averaged			**74.21**

Table 8 Final Optimal Parameters for System Training

Parameters	**Description**
Input values	Binary
Number of input neurons	2500
Number of hidden neurons	500
Number of output neurons	3
Learning Rate	0.8
Activation Function	Sigmoid
Stopping Condition	Epoch < 10000 & MSE < 0.0001

6 Conclusions

The purpose of this research is to develop a prototype that is able to recognize paddy disease and give recommended solutions based on the type and stages of the paddy disease. This is to help paddy farmers and also paddy researches to be able to take quickly take action when faced with this problem. Even with all of the challenges faced in this project, the strengths, limitations and recommendations are valuable for those who are going for the same type of problem domain or even for those who are implementing any type of image processing problems. The problem of external forces such as lighting, background problems and also noises are the problems that need to be tackled. The project also opens up numerous future works to be done and improved in the field of image processing and also paddy plant disease recognition.

Acknowledgement:

The authors would like to thank Malaysian Agriculture Research and Development Institute (MARDI) for their cooperation in making this project successful.

References:

[1] Anim, H. *Penyakit Karah Padi*. from http://animhosnan.blogspot.com/2012/03/penyakit-karah-padi.html.

[2] Anthonys. G., & Wickramarachchi. N. (2009). *An Image Recognition System For Crop Disease Identification of Paddy Fields In Sri Lanka*. Paper presented at the Fourth International Conference on Industrial and Information Systems, ICIIS December 2009.

[3] Husin. Z., Shakaff. A. Y. M., et al. (2012). *Embedded Potable Device For Herb Leaves Recognition Using Image Processing Techniques And Neural Network Algorithm.*

[4] Inside Malaysia (2012). *Agriculture & Aquaculture*

[5] IRRI (1996). *Standard Evaluating System* (SES) *for Rice (4th Edition).*

[6] IRRI (2014). *Rice Knowledge Bank*Jack, A. (2014). Malaysian Agriculture of Research Development Institute.

[7] Kamarudin. S, Noraidah. S, et al. (2013). *A Framework Of An Expert System For Crop Pest And Disease Management*.

[8] Kurniawati, N.N., Abdullah, S, et.al. (2009). *Texture analysis for diagnosing paddy disease*. Paper presented at the International Conference on Electrical engineering and Informatics, August 2009.

[9] MalaysiaDigest (2012). *Growth Domestic Product (GDP) Malaysia 2012*.

[10] Nooro, A. N. (2014). *Development of Shrubs Recognition System using Backpropagation Neural Netwok.*

[11] Nursuriati, J., & Khalil A. (2014). *Practical Digital Image Processing with MATLAB.*

[12] Saad, A. and H. Habibuddin (2010). *Pathotypes and Virulence of Xanthomonas Oryzae causing Bacterial Blight Disease of Rice in Peninsular Malaysia* (Patotip dan virulens Xanthomonas oryzae yang menyebabkan penyakit hawar bakteria padi di Semenanjung Malaysia). J. Trop. Agric. and Fd. Sc. 38(2) (2010): 257–266.

[13] Sankaran. S., Mishra. A., Ehsani. R & Davis. C. (2010). *A review of advanced techniques for detecting plant diseases*. c (72): 1-13.

[14] TradingEconomics (2014). *Employment in Agriculture (% of total employment) in Malaysia*.

ISBN: 978-1-61804-302-3