INTELLIGENT PADDY RICE COLOR RECOGNITION SUITABLE FOR HARVESTING

Athirah A.Rahim
2004257182
Faculty of Electrical Engineering
Universiti Teknologi Mara
40450 Shah Alam
chilla1718@yahoo.com

Abstract

This paper presents an automatic recognition of paddy rice color using RGB color extraction. In this work, five sets of paddy rice images from paddy field at Kampung Tua, Semanggol Perak are digitally captured at ICS (Image Capturing Studio) room. The identified regions of interest (ROI) of these paddy's images are processed to quantify the reflectance indices in RGB color model. Paddy rice images are then processed to produce the dominant RGB pixel indices in the primary color model. These reflectance indices gained under standard and controlled environment are then used to design a ANN diagnosis model for paddy rice using MATLAB software. The optimized model is evaluated and validated through analysis of the performance indicators regularly applied in classification models. From the findings, this work has shown that the best model has produced percentage accuracy of 88.75%, 92% specificity and 85.5% sensitivity when measured at 0.1 threshold with a balanced percentage rate of training dataset.

Keyword: Paddy rice color, RGB color extraction, MATLAB programming.

1.0 INTRODUCTION

Rice is the dominant staple food of Asia, accounting for more than 70% of caloric intake in some countries. Therefore, paddy research can play a key role in improving quality of the rice produced. In addition, paddy research contributes to improve biotechnology system through several pathways, and these contributions benefit both producers and consumers.

Conventionally, maturity of paddy rice is decided by counting the numbers of days after they are planted. Depending on their varieties, usually harvesting would take into effect between 105 – 125 days where generally 80% of the paddy rice is ripe. "MR 220 padimas" is a type of paddy that has maturity as early as 110 days and is recommended by MARDI [1]. For this type of paddy, the plant starts to produce paddy rice after 60 days from seedling. Within the period of 45 days later, the paddy rice becomes ripe. However in large-scale production, period of paddy maturity is inconsistent due to the different cultivating time. Since paddy rice presentation can also be presented in terms of digital images, therefore they can be processed and measured to produce important quantitative features information.

This research proposed that the above selected features information to be used in designing an intelligent decision model for the accurate color features of the matured paddy rice for harvesting. The front end of the model will utilize color index from the paddy rice using the latest image processing technique. Quantified parameters that can represent the best color features are the mean and standard deviation of the paddy RGB image samples. The information are then being used to train a 3-layer Artificial Neural Network (ANN) using back propagation algorithm; Lavenberg Marquard. Optimization of these experimented models are evaluated or validated by observing the percentage of classification rate.

2.0 METHODOLOGY

2.1 Paddy Rice Samples

The paddy samples used in this study were collected from Kampung Tua Semanggol Perak. Five sets of samples paddy images representing the interval of ten

days and five days for class 1 and 2 and class 3, 4 and 5 respectively were used for this study. It should be noted here these sets are categorized as controlled data as the images were captured under controlled environment. There are 1400 samples of data have been collected which are 1000 samples for training and 400 samples for testing. The data have been recorded in mean values of red, green and blue (RGB) color model. Figure 1 shows the cluster of the paddy images.

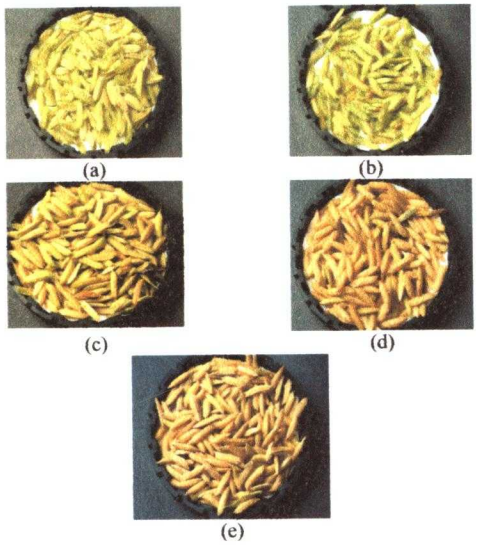


Figure 1: Types of Paddy Image depend to interval time; (a) cluster1; 80-89 days (b) cluster2; 90-99 days (c) cluster3; 100-104 days (d) cluster4; 104-105 days (e) cluster5; 110 days.

2.4 Instruments and Measuring Procedures

The Red, Green, Blue (RGB) component color images were acquired using FinePix 6900 Zoom (FujiFilm) digital camera, with pixel resolution of 1280x960 and save in JPEG image. The image capturing process had been done under standard and control environment in Image Capturing Studio Room (ICS Room) at Advanced Signal Processing (ASP) Research Lab, Faculty of Electrical Engineering, UiTM Shah Alam. The camera was placed at a distance of one foot directly above the paddy samples with 55° of light setting angle. The lighting used for capturing image is sportlight Digicolor K-250C, AC 170~240V/50Hz. The light intensity was controlled by using Heavy Duty Light Meter (Model 407026) and Heavy Duty Data Logger (Model 380340) having mean lux of 2682 $\pm$ 43.

Figure2: The setting on capturing devices (a, b and c)

The process of classification the paddy sample can be described in Figure 3:

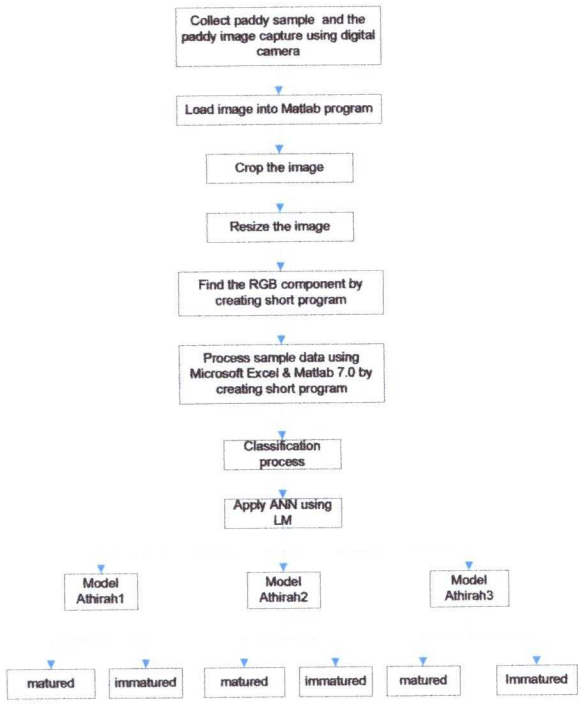


Figure 3: Flow chart of the Intelligent Paddy Rice Color Recognition

2.3 Digital Image Processing

The process started with each image from the stored database was being filtered. There were artifacts such as small white ellipse lines or dots within the image which was recommended to remove using median filter technique [2].These artifacts can be considered as impulsive noise and may thus be reduced using a median filter given by:

$$P_{med}(m,,n) = median\{P(m-k,n-1)\,|\, -\frac{N_{med}-1}{2} \le k,$$

$$l \le \frac{N_{med}-1}{2} \wedge 1 \le m-k \le m \wedge 1 \le n-1 \le N\ \} \qquad (1)$$

where N_{med} is odd[2] and indicates the size of the two dimensional median filter. P represents all the three color components. Note that only square median filter kernel was considered.

After the filtering process, region of interest (ROI) which includes a sample of paddy images were selected. Each image was carefully studied and observed. After the regions of image have been identified, they were cropped out sequentially with the first cluster image and followed by the other cluster image sample. All samples were then been resized to a dimension of 50 x 50 pixel area refer to previous literature review [3]. Furthermore, standardizing all the sample area was necessary for easier computation during quantification of color extraction.

2.4 ANN Model

In this research, the main objective is to identify the suitable paddy for harvesting based on the extracted RGB dominant pixel gradation which are the mean indices gathered from sample images. Learning rule is defined as a procedure for modifying the weights and biases of a network. (This procedure may also be referred to as a training algorithm.) The learning rule is applied to train the network to perform some particular task. Learning rules is divided into two broad categories: supervised learning, and unsupervised learning. This research use unsupervised learning which is provided with the weights and biases that are modified in response to network inputs only. There are no target outputs available. Most of these algorithms perform clustering operations. They categorize the input patterns into a finite number of classes. This is especially useful in such applications as vector quantization [4].

The type of network that best fits in the diagnosis application is the multilayered perceptron (MLP) network with one hidden layer based on the fact that it has been widely satisfactorily applied by many researchers [5]. The network is shown in Figure 4 and the output can be expressed by the following equations:

$$\hat{y}_i = \sum_{j=1}^{t} w_{ij}^2 f\left\{\sum_{k=1}^{s} w_{jk}^1 x_k^0 + b_j\right\} \tag{3}$$

where $i = 1$ and w_{ij}^2, w_{jk}^1, b_j denotes the adaptive variables to be optimized and their values are changed many times during the network training process.

The $f(\bullet)$, is the sigmoidal linear activation function which was used to map the sum of the weighted inputs to the output of a neuron in the hidden layer. The network was trained using Lavenberg Marquard algorithm.

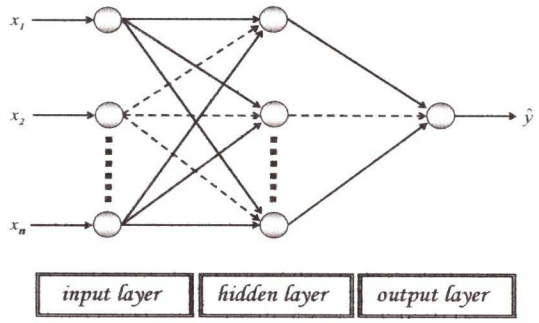


Figure 4: Figure caption.

Sensitivity and specificity are commonly used terms that generally describe the accuracy of a test. Sensitivity is a measure of the ratio or percentage of 'true' positive (TP) and a positive diagnostic test result ($D+$). It represents the actual percentage of a 'true' (positive) samples paddy image realized by a positive test result and is also known as true positive rate (TPR), defined as:

$$\text{Sensitivity:} \qquad TPR = \frac{TP}{D+} \tag{5}$$

Specificity measures the ratio or percentage of 'true' negative (TN) and with a negative diagnostic test result ($D-$). It is actually represents the actual percentage of a 'false' (negative) samples condition realized by a negative diagnostic test. Specificity is also termed as true negative rate (TNR) and is given as:

$$\text{Specificity:} \qquad TNR = \frac{TN}{D-} \tag{6}$$

The percentage for diagnostic accuracy refers to the percentage of samples that have been correctly classified or diagnosed, and have output values within the predefined threshold range for the respective output level. It can be derived as:

$$DA = \frac{\sum_{i=1}^{N} c_i}{N} x100\% \tag{7}$$

Variable c_i serves as counter for the proposed ANN model output, y at sample i. c_i is defined as:

$$c_i = \begin{cases} 1, & 1-thr \le y \le 1+thr \quad for \text{true} \\ 1, & -1-thr \le y \le -1+thr \quad for \text{false} \\ 0, & elsewhere \end{cases} \tag{8}$$

During the training phase process, the ANN model was optimized to find optimum value or setting for the number of neurons in the hidden layer. Selection of the best numbers of neurons was based on the performance evaluation of the model through the sum-squared error (*SSE*) analysis as well the diagnostic accuracy (*DA*). For simplicity, threshold for the output logic levels was fixed to $\pm$ 1 for each model in this initial work. At a later stage, the most appropriate threshold level would be decided by analyzing the minimum *Euclidean Distance* (*ED*) values form the receiver operating characteristic (ROC) plot and followed by the clustering plots [5].

2.5 Lavenberg Marquard

The Levenberg-Marquardt algorithm was designed to approach second-order training speed without having to compute the Hessian matrix[4]. When the performance function has the form of a sum of squares (as is typical in training feed forward networks), then the Hessian matrix can be approximated as,

$$H = J^T J \tag{9}$$

And the gradient can be computed as;

$$g = J^T e \tag{10}$$

Where is the Jacobian matrix, which contains first derivatives of the network errors with respect to the weights and biases, and **e** is a vector of network errors. The Jacobian matrix can be computed through a standard back propagation technique that is much less complex than computing the Hessian matrix [4]. The Levenberg-Marquardt algorithm uses this approximation to the Hessian matrix in the following Newton-like update:

$$x_{k+1} = x_k J^T - \left[J + mI\right]^{-1} J^T e \tag{11}$$

The LM algorithm is also used to train the ANN. The LM algorithm attempts to solve a nonlinear least square minimization problem in the form of (12):

$$f(x) = \frac{1}{2}\|r(x)\|^2 \tag{12}$$

Where r is the residual vector (13);

$$r(x) = \left[r_1(x), r_2(x), \ldots, r_m(x)\right] \tag{13}$$

The LM rule uses complementary properties of the Vanilla Gradient Descent (VGD) and the Gauss-Newton (GN) rules [6]. The LM rule is given by (14);

$$x_{i+1} = x_i - \left(H - \lambda diag\left[H\right]\right)^{-1} \nabla f(x_i) \tag{14}$$

2.6 Experimental Design

The experiments were designed to test the effectiveness of the input features in discriminating the matured (positive data) from the immature (negative data) paddy images. This experiment used a total of 1400 samples where 1000 samples were for training and the rest was for testing. In order to generate the best classification network possible, the size of the training set for positive data and negative data should be maximized or minimized. However, various size of training data sets were preferred in this study in order to provide a wider spectrum of results, and hence to allow for observation of trends in the data. The proposed sets of training data are shown in Table 1 as below. All of the three models have the same three inputs representing mean value of RGB. The network was trained and tested repeatedly with different random subsets of these training input data.

Table 1: Sets of training data

Model	Training data		Testing data	
	+ve: -ve (samples)	*+ve: -ve (%)*	*+ve: -ve (samples)*	*+ve: -ve (%)*
Athirah1	600:400	60:40	200:200	50:50
Athirah2	400:400	50:50	200:200	50:50
Athirah3	700:300	70:30	200:200	50:50

3.0 RESULTS AND DISCUSSION

Figure 5 shows an example of SSE performance with respect to epoch iterations during training of best hidden layer size. Shown only are selected number of neurons; 3, 5, 7, 11 and 13 for Athirah1 model where their respective SSE are observed to be converging as it approaches to 1000 epochs. Thus, implying that training of the model is successful for each experiment.

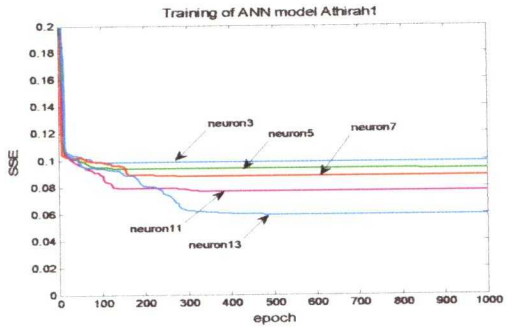


Figure 5: Epoch performance with respect to the hidden layer size.

Figure 6 shows the performance plot for Athirah1 model. It can be observed that highest accuracy is observed when the hidden layer size has 5 neurons. The accuracy achievement reaches 96.5% at this size. Increasing number of neurons does not improve the performance as the accuracy decreases. This is due to the network model has reached overfitting where the neurons were trained to recognize the training data very precisely.

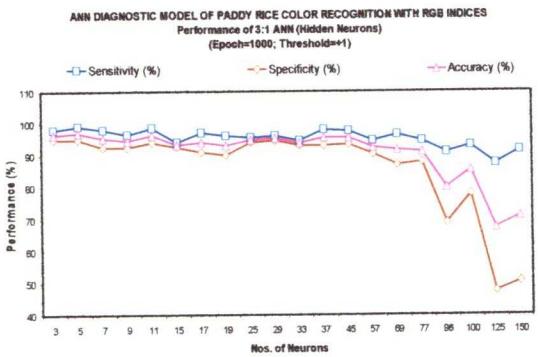


Figure 6: ANN model Athirah1 showing the training and testing results of neurons. The best value obtained is 5 neurons.

Figure 7 shows the performance plot of 5 best selected neuron; 3, 5, 7, 9 and 11 for model Athirah2. From the plot, it is shown that SSE performance converges for each model as it approaches to 1000 epochs. This implies that the training of the models is successful for each experiment.

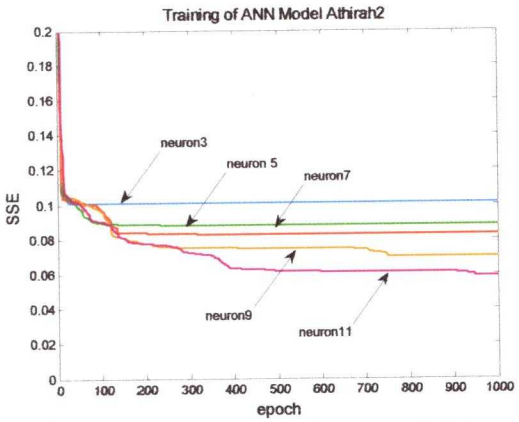


Figure 7: Epoch performance with respect to the hidden layer size.

Figure 8 shows the performance plot for model Athirah2. The accuracy achievement reaches 88.5% when the hidden layer size has 5 neurons.

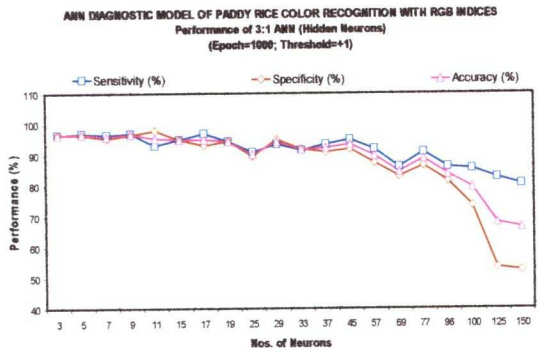


Figure 8: ANN model Athirah2 showing the training and testing results of neurons. The best value obtained is 5 neurons.

Figure 9 shows the performance of 5 best selected neuron; 5, 7, 11, 15 and 17 for Athirah3 model. SSE performances are detected to be converged for each model as it approaches to 1000 epoch. Thus the training is successful.

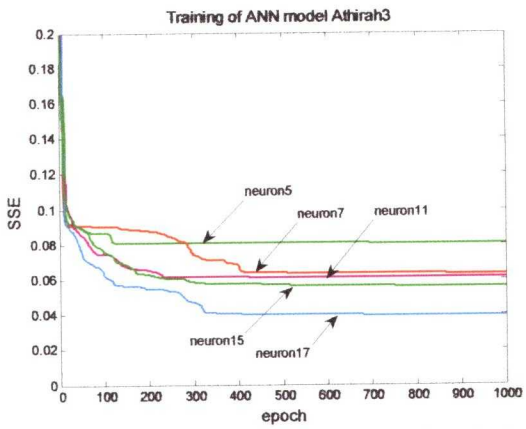


Figure 9: Epoch performance with respect to the hidden layer size.

The accuracy achievement for Figure 10 (Model Athirah3) reaches 95.75% when the hidden layer size has 3 neurons.

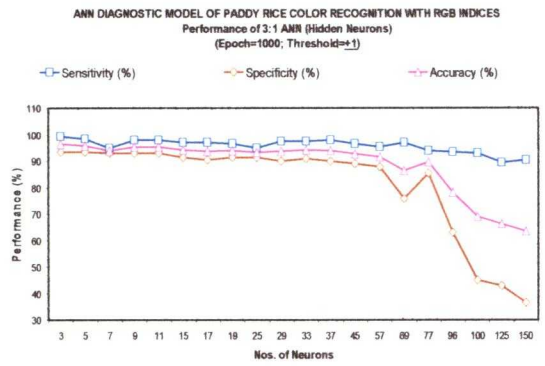
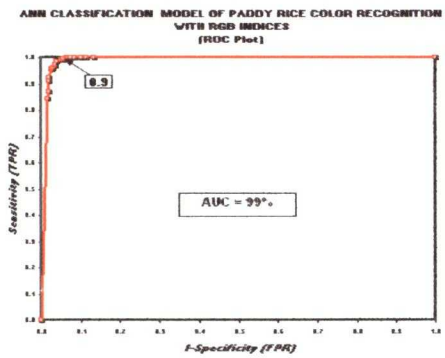


Figure 10: ANN model Athirah3 showing the training and testing results of neurons. The best value obtained is 3 neurons.

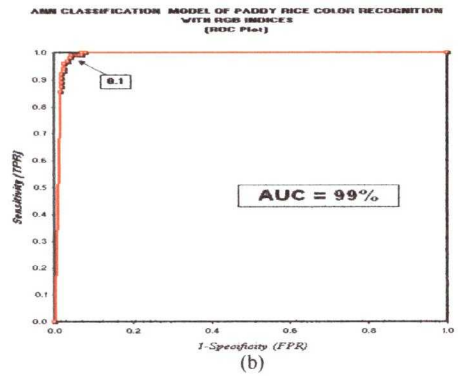
Figure 11 shows the ROC plots for all models where all of the three models have good percentage of accuracies, indicated by:

a. Both curves leaning very closely to the true positive rate (*TPR*) and false positive rate (*FPR=1-TNR*) axis.
b. The areas under the curve are close to 100% indicating all models have good specificity and sensitivity scores [6].

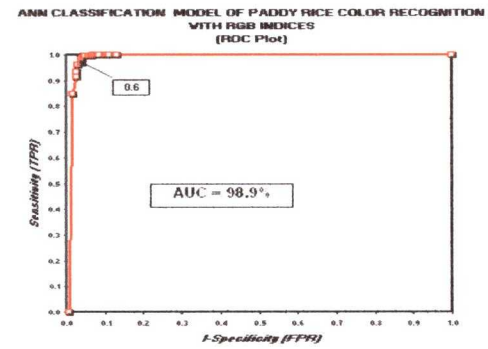
Also from the plots, the best threshold identified to have the shortest Euclidean distance from the ideal point at the top-left hand corner is 0.9, 0.1 and 0.6 for each respective model Athirah1, Athirah2 and Athirah3. A scatter plot is drawn to display the clustering of simulated output between the expected target levels. The scatter plot for all three models is shown in Figure 12. It can be observed that majority of the outputs are concentrated at both target levels implying high accuracy.



(a)

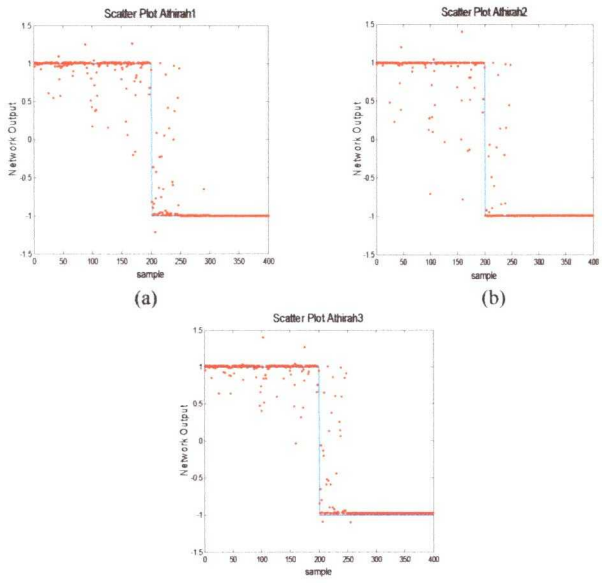


(b)



(c)

Figure 11: ROC plot comparison between three models; (a) model Athirah1; AUC 99% (bj) model Athirah2; AUC 99% (c) model Athirah3; AUC 98.9%.



(a) (b)

(c)

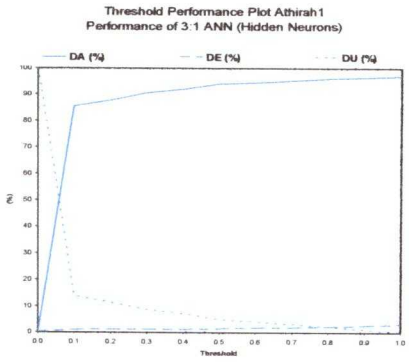
Figure 12: Scatter plot the three models; (a) model Athirah1 (b) model Athirah2 (c) models Athirah3. The solid line represent the expected output, the dots represent the network-simulated output.

A threshold is defined to examine the clustering of network outputs at different thresholds. The threshold starts at 0 and is incrementally increased. The unsure region shrinks as the threshold increases. At lower

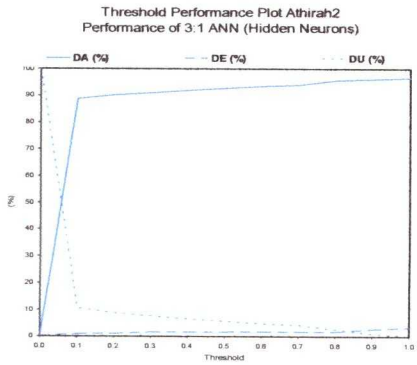
threshold levels, only outputs that reside much closed to the expected output are considered to be correct. Thus, as the threshold increases, DU will decrease while DA and DE will typically increase. A good classifier will have these properties [6]:

a. DA should be high across the threshold, with a small increase in DA as the threshold is increased. This indicates that the network can still accurately classify cases at low thresholds due to very good clustering near the expected value.
b. DU is low across the threshold. A good classifier will exhibit a small decrease in DU when the threshold is increased. This property indicates that there are only a few cases that cannot be satisfactorily classified as matured or immature paddy as the threshold increased.
c. DE is low across the threshold. A good classifier will exhibit a small increase in DE when the threshold is increased. This property indicates that there are only a few cases that are incorrectly classified by the network.

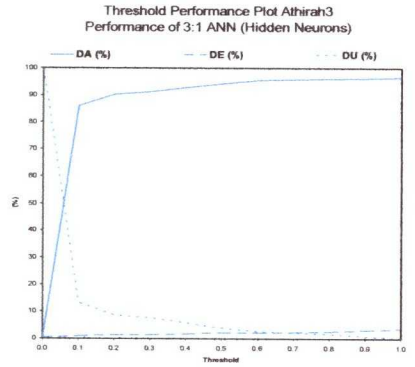
The effect of threshold adjustment is shown in Figure 13. By considering at *thr=0.1* because of high clustering, Athirah2's DU (10%) have lower diagnostic rates and higher DA (88.75%) than Athirah1 (DU = 14% and DA = 85.5%) and Athirah3 (DU = 13% and DA = 86%). Each systems exhibit low DE and DU which is below 5% at these thresholds. Thus, this threshold is recommended to be used for classification. With respect to this threshold, Athirah2 has shown better performance than Athirah1 and Athirah3 due to its good diagnostic rates.



(a)



(b)



(c)

Figure 13: Graph shows the threshold plot for all three model; (a) model Athirah1; threshold 0.9 (b) model Athirah2; threshold 0.1 (c) model Athirah3; threshold 0.6

Table 2 shows the summarization of the necessary performance indicator for all three models. From the table, Athirah2 has better accuracy, sensitivity and specificity amongst the other two models. However, the network size is slightly higher than Athirah3.

Table 2: Comparison between ANN models for this experiment.

Spectrum Model Of Testing Dataset			
Model	**Athirah 1**	**Athirah 2**	**Athirah 3**
TRAIN DATASET (%+vedata:%-ve data)	60:40	50:50	70:30
ANN (Input:Hidden:Output)	3:5:1	3:5:1	3:3:1
ANN (No. of connection)	24	24	16
Threshold	0.1	0.1	0.1
Sensitivity (%)	84.50	85.50	85.00
Specificity (%)	86.50	92.00	87.00
DA (%)	85.50	88.75	86.00
AUC	99%	99%	98.9%

4.0 CONCLUSION AND FUTURE RECOMMENDATION

4.1 Conclusion

The ANN is a suitable technique to identify the suitable paddy rice for harvesting based on the dominant RGB color pixel (mean). The ANN consists of a hidden layer, each locally connected to an input feature. Three models with different spectrum model were analyzed and compared to get the desired result. Training was performed using the LM algorithm with 1000 epochs. Performances of the methods were displayed using performance indicator table and graph.

The result have shown that the best model with excellent abilities was Athirah2 since it recorded the highest percentage of accuracy (88.75%), high specificity (92%) and sensitivity (85.5%) at 0.1 threshold with the rate of training dataset is 50:50. This model Athirah2 can be concluded as the best model which has equal size for training positive data and negative data. Even though the number of training data is different, model Athirah1, Athirah2 and Athirah3 consumed small size of network connection, thus creating lower cost for production.

4.2 Future Recommendation

ANN theory and practice suggest that, in a diagnostic application, the network should be trained with a balanced mixture of inputs from each testing and training data [7]. For future recommendation, instead of using the different subsets of training dataset only, it also recommended to use different subset of training and testing datasets both. In order to generate the best classification network possible, the size of the training set should be maximized. On the other hand to improve the confidence level in the results as an estimate of future performance, the size of test set should be maximized. This suggests an equal size for training and testing sets. The other recommendation is by using the real pixel value representing the cropped image and further reduces the input size using PCA technique. However, to implement this strategy, the number of dataset must be greater than input size.

5.0 ACKNOWLEDGEMENT

Sincerely, I would like to thanks my project supervisor, Prof Madya Dr Hadzli Hashim for his valuable advice, ideas and critical guidance. A lot of thankful to my parents especially my father, Mr. A.Rahim Kasman who has supplied all of the paddy rice samples, and to all my friends for their unfaltering commitment, support and immense faith during this research. Always a joy to work with, all of you makes the difficult task of authoring profoundly enjoyable.

6.0 REFERENCES

[1] R&D, "MR 219/ MR220," in *Malaysian Agricultural Research and Development Institute*. R&D Technology and Technical Services in Tropical Agricultural and Food.

[2] H. Hashim, R.Jarmin, and R.Jailani, "Plaque Psoriasis Diagnosis Model with Dominant Pixel Gradation from Primary Color Space," in *ASP Research Group, Faculty of Electrical Engineering*. Shah Alam, Selangor: Universiti Teknologi MARA.

[3] J. Paliwal, N. S. Shashidar, and D. S. Jayas, "Grain Kernel Identification using Kernel Signature."

[4] MATLAB, "MATLAB Programming Lavenberg Marquard," in *Help Lavenberg Marquard*, 7.0, Ed.

[5] S. S. A. Razak and H. Hashim, "Automated Plaque Diagnosis Model Utilising Gaussian RGB Parameters " in *ASP Research Group, Faculty Of Electricl Engineering*. Shah Alam, Selangor: Universiti Teknologi MARA.

[6] I. M. Yassin and M. N.Taib, "Particle Swarm Optimization as Network Optimizer for Neural Network Based Face Detector," in *IEEE member*.

[7] F. Ercal, A. Chawla, W. V. Stoecker, H.-C. Lee, and R. H. Moss, "Neural Network Diagnosis of Malignant Melanoma From Color Images," in *Senior Member IEEE*.