How to cite this paper:

Mohd Zulhilmi Ab Jamil, Sofianita Mutalib, Shuzlina Abdul Rahman, & Zalilah Abd Aziz. (2017). Paddy weed leaf classification using neuro-fuzzy system in Zulikha, J. & N. H. Zakaria (Eds.), Proceedings of the 6th International Conference on Computing & Informatics (pp 601-606). Sintok: School of Computing.

PADDY WEED LEAF CLASSIFICATION USING NEURO-FUZZY SYSTEM

Mohd Zulhilmi Ab Jamil[1], Sofianita Mutalib[2], Shuzlina Abdul Rahman[3], Zalilah Abd Aziz[4]

[1]*Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA, 40450 Shah Alam, Malaysia, mz.hilmi@gmail.com*

[2] *Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA, 40450 Shah Alam, Malaysia, sofi@tmsk.uitm.edu.my*

[3] *Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA, 40450 Shah Alam, Malaysia, shuzlina@tmsk.uitm.edu.my*

[4] *Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA, 40450 Shah Alam, Malaysia, zalilah@tmsk.uitm.edu.my*

ABSTRACT. The productivity of paddy rice in Malaysia has improved as the technology advances. However, several concerns need to be considered to maintain the productivity. The growth of paddy weeds in the fields is one of the concerns that need to be taken care of as it resulted to the decrease of the paddy yield. This paper presents an integrated method for classifying paddy weeds through the shape of their leaves by applying neuro-fuzzy techniques. The focus weed types are Sphenoclea zeylanica, Ludwigia hyssopifolia and Echinochloa crus-galli. The developed e-prototype is able to classify the paddy weeds with 83.78% accuracy. Hopefully the findings in this study may assist the farmers and researchers in increasing their paddy yield.

Keywords: Fuzzy Logic, Neural Network, Neuro-Fuzzy, Paddy Weeds.

INTRODUCTION

Rice is the highest food grain consumed by nearly 50% of the world's population (Kwak, et al., 2008). In Malaysia, paddy is important in terms of food security and economy. Until now, many researches have been conducted by several agencies like Malaysian Agricultural Research and Development Institute (MARDI) to increase the paddy productivity in Malaysia such by introducing the new variety of rice and aerobic paddy which use less water.

However, there is a wild threat in the paddy field that is the growth of weeds which in turn will increase paddy yield loss. Almost ten percent of the yearly rice yield loss worldwide is caused by the paddy weeds (Rabbani et al., 2011). This is because these weeds can compete with paddy for nutrients, water and sunlight (Hakim et al., 2013).

Therefore, it is important for the paddy farmers to recognize the paddy weeds at their paddy fields to overcome them effectively. The proposed prototype is expected to help the farmer in recognizing paddy weeds using captured images. This paper is organized into five sections. It begins with the introduction and some related works on paddy weeds. Subsequently, the methodology and the experimental design of the works were described. This is then followed by the results and conclusion and finally, ended with the conclusion.

PADDY WEEDS

There are three types of weeds which are sedges, grasses and broadleaved weeds. Hakim et. al (2013) stated that the paddy loss caused by the weeds are different from one place to another and it depends on the predominant flora and the control techniques used by the paddy farmers. In Malaysia, a lot of researches have been done to improve the farmers weed control management such as developing a better breed of rice paddy and enhancing the current herbicides (Rahman et al. (2012) and Juraimi et al. (2013)).

There are several locations in Malaysia that have been identified with the highest growth rate in the paddy weeds in the fields, which are in Tanjong Karang, Selangor (Hakim et al., 2011), Kuala Muda, Kedah (Mansor et al., 2012) and Seberang Perak, Perak (Hakim et al., 2013). Based on these studies, it is shown that Echinochloa crus-galli, Ludwigia hyssopifolia and Sphenoclea zeylanica are the most common paddy weeds detected in Malaysia paddy fields.

Echinochloa crus-galli

According to Man et al. (2015), Echinochloa crus-galli or “Rumput sambau” is from poaceae family. The leaf length is between 8 to 60 centimetres. The leaf is glaborous or hairless. It can be controlled chemically by using molinate or quinclorac.

Ludwigia hyssopifolia

Man et. al (2015) stated that Ludwigia hyssopifolia is from onagracae family. In Malay, it is called “maman pasir”. The leaf is 2 to 10 centimetres long and is lance-shaped. To manage the weed, the chemicals used are glufosinate ammonium or glyphosate if it is on the edge of the paddy field. In the field, 2,4-D or cyclosulfamuron is recommended.

Sphenoclea zeylanica

Sphenoclea zeylanica or “cabai kera” in Malay comes from spenochleaceae family (Man et al., 2015). Its leaf can go up to 10 centimetres long and it is usually oblong or lance-shaped. To control it, bensuron-methyl is recommended to use in the fields.

METHODS

This section will discuss the key approaches by first; taking the raw image (data) and perform several image pre-processing techniques such as cropping and rescaling. Next, the image is then being processed through conversion into grayscale and binary form to extract the features of the leaf. Finally, neuro-fuzzy system will determine the type of the paddy weed based on the features of the leaf.

Data for this study was obtained from paddy fields in Seberang Perai, Pulau Pinang and Sungai Besar, Selangor. Another resource is from Pl@ntNet, a repository of images of the paddy weeds. Most of the data obtained was in the form of Portable Network Graphics (PNG) format due to its lossless data characteristic. 248 samples were collected for the experiment. Table 1 shows the number of samples collected.

Table 1. Number of samples collected

Plant weed	No. of samples
Echinochloa crus-galli	48
Ludwigia hyssopifolia	100
Sphenoclea zeylanica	100

Image Processing

Several image processing techniques were used before the feature extraction. The techniques are orientation, cropping and scaling conversion in which the images are converted into grayscale and black and white colour. It is important so that the features of the leaf can be extracted effectively. Figure 1 shows the flow of the image processing techniques involved while Table 2 shows the ratio, R which is equal to the length of the image/width of the image and the rescaling process.

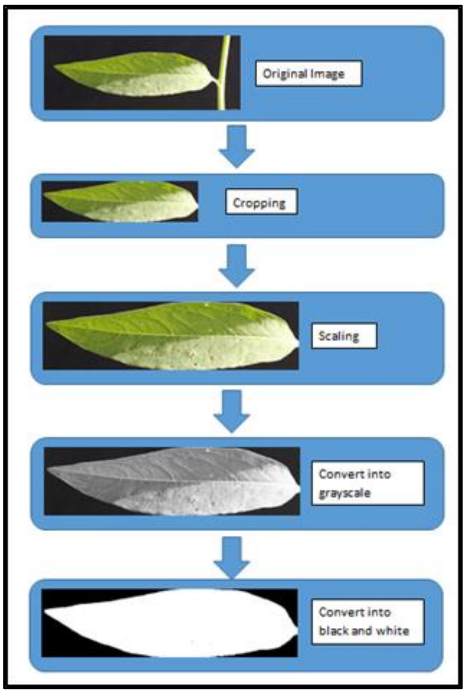

Figure 1. Image processing flow

Table 2. Ratio and scaling dimension

Ratio, R	Dimension (l x w)
R <= 1.4	300 x 300
1.4 < R <= 2	300 x 200
2 < R <= 2.4	200 x 140
2.4 < R <= 3	200 x 100
3 < R <= 3.5	300 x 65
R > 3.5	400 x 45

Feature Extraction

There are four fuzzy variables for this study which are length (L), width (W), perimeter (P) and area (A) that resembles the shape of the leaf. The samples are analyzed by using discretization to identify the membership functions of each variable. The membership functions of the variables are shown in Table 3. After the discretization process, a decision table is made to generate rules. The rules are based through the observation from the discretized data. The rules are then set to 12, 24, 36 and 48 prior to testing the accuracy of the set of rules.

Table 3. Membership functions of the fuzzy variables

Fuzzy value / Fuzzy variables	Short (L & W) Small (P & A)	Medium	Long (L & W) Large (P & A)
Length (L)	0 – 200	100 – 300	200 – 400
Width (W)	0 – 150	100 – 250	200 – 300
Perimeter (P)	0 – 600	300 – 900	600 – 1200
Area (A)	0 - 20000	10000 - 30000	25000 - 50000

Architecture Design

Once the rules have been generated, the next process was to design the architecture of the neuro-fuzzy system. Neuro-fuzzy system integrates artificial neural network (ANN) and fuzzy logic system. Generally neuro-fuzzy system has input, output and three hidden layers which represent the membership functions and fuzzy rules (Borah et al., 2015). Therefore, neuro-fuzzy system incorporates the computation and learning capabilities of ANN with the human-like knowledge and representation from the fuzzy logic system. Figure 2 shows the neuro-fuzzy architecture that was designed for this study.

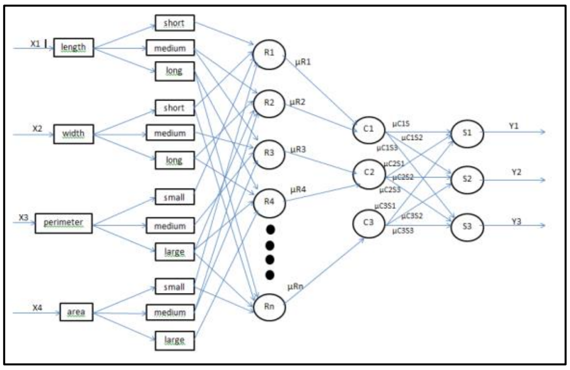

Figure 2. Neuro-fuzzy system architecture

Experimental Setup

The data was split into two sets, training and testing set. Different split ratios used are 60:40, 70:30 and 80:20. For the parameters, the learning rate was set from 0.5 to 1. The number of epochs was ranged from 200,000 to 1,000,000 with 200,000 intervals.

The training process started by receiving the measurements of image data in the excel file format (.CSV). Then, the system was retrained using different set of rules for evaluation. After that, the system was tested with different set of images to determine the accuracy of the system.

RESULTS AND DISCUSSION

During this phase, the prototype was tested for its correctness in classifying the paddy weeds. The testing data was in the form of excel data file (.CSV) which contains the measurements of the leaf. The accuracy of the data was determined after the testing phase was completed. The accuracy of the system is calculated using the formula in Eq. (1):

$$Accuracy\ result = (Number\ of\ correct\ samples)\ /\ (Number\ of\ total\ samples)\ *100\% \quad (1)$$

Before finding the optimal parameters for the system, the set of rules was evaluated first. It is based on the average training and testing accuracy of each number of rules. Based on Table 4, it is shown that higher set of rules produce lower mean squared error and the ratio for three datasets are given in the same table.

Table 4. Average mean squared error on different datasets and number of rules

Dataset	Average Mean Squared Error			
	12 Rules	24 Rules	36 Rules	48 Rules
1 (60:40)	0.005544	0.004809	0.003952	0.002181
2 (70:30)	0.004887	0.004377	0.003499	0.002557
3 (80:20)	0.004921	0.004131	0.003735	0.002521

By using 48 rules, the next experiment is to identify the optimal learning rate for each datasets. In this experiment, it is shown that Dataset 1 (60:40) has the lowest average mean squared error when the learning rate is 0.7. However, Dataset 2 (70:30) and 3 (80:20) has the lowest average MSE when the learning rate is 0.9. Table 5 shows the results of the all the datasets on different learning rates.

Table 5. Average MSE on different learning rates for all datasets

Learning rate	Average MSE		
	Dataset 1	Dataset 2	Dataset 3
0.5	0.002506	0.002912	0.003048
0.6	0.002364	0.002488	0.00262
0.7	0.001928	0.002516	0.002378
0.8	0.002366	0.002548	0.002155
0.9	0.001938	0.00221	0.002064
1.0	0.001986	0.002668	0.002372

Besides average mean squared error, testing accuracy was calculated as well. For Dataset 1, the best learning rate is 0.5 meanwhile Dataset 2 and 3 are 0.6 and 0.8 respectively. Figure 3 shows the graph of the average accuracy based on learning rate.

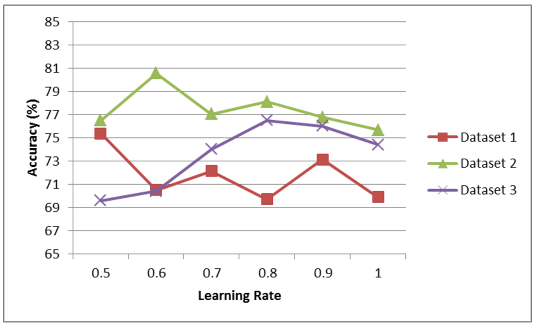

Figure 3. Average accuracy on learning rate

Using the optimal learning rate for each datasets as mentioned above, the number of epochs is tested for better accuracy. The number of epochs is ranged between 200,000 and 1,000,000 epochs. Based on the experiment, Dataset 1 has the highest accuracy using 200,000 epochs while Dataset 2 and 3 iterated 1,000,000 epochs to gain the highest accuracy. Table 6 shows the optimal parameters for each datasets and their results.

Table 6. Optimal parameters for each datasets

Dataset	1	2	3
Learning Rate	0.5	0.6	0.8
Number of epochs	200,000	1,000,000	1,000,000
Mean Squared Error	0.0036	0.00173	0.0016
Accuracy (%)	78.79	83.78	82

CONCLUSIONS

This study has done analysis on constructing fuzzy sets and neuro-fuzzy architecture in the case of interest. Later, we developed a prototype that can classify the paddy weeds through the shape of the leaves by applying the method. With an accuracy of 83.78%, this prototype can be considered reliable and hopefully it can help paddy farmers to take quick actions when dealing with paddy weeds in their paddy fields. However, there are several limitations that need to be tackled such as the environmental settings of the image taken and future enhancement can be done based on these limitations.

ACKNOWLEDGEMENT

The authors would like to express the gratitude to all related parties, experts from MARDI, Kepala Batas, REI 19/2015 for the research support and also to IRMI UiTM, Shah Alam.

REFERENCES

Borah, T. R., Sarma, K. K., & Talukdar, P. H. (2015). Retina Recognition System using Adaptive Neuro Fuzzy Inference System. Paper presented at the IEEE International Conference on Computer, Communication and Control (IC4-2015).

Hakim, M. A., Juraimi, A. S., Hanafi, M. M., Selamat, A., Razi Ismail, M., & Anwar, M. P. (2011). Weed flora in rice fields of Tanjong Karang coastal area in Malaysia. Journal of Food, Agriculture & Environment, 9(1), 694-699.

Hakim, M. A., Juraimi, A. S., Razi Ismail, M., Hanafi, M. M., & Selamat, A. (2013). Survey on weed diversity in coastal rice fields of Seberang Perak in Peninsular Malaysia. The Journal of Animal & Plant Sciences, 23(2), 534-542.

Juraimi, A. S., Uddin, M. K., Anwar, M. P., Mohamed, M. T. M., Ismail, M. R., & Man, A. (2013). Sustainable weed management in direct seeded rice culture: A review. Australian Journal of Crop Science, 7(7), 989-1002.

Kwak, T. S., Whang, H. G., & Chung, C. S. (2008). Importance of Rice Genetic Resources and Current Status of Rice Breeding and Cultivation in Korea. Journal of the Faculty of Agriculture Shinshu University, 44(1), 35-37.

Man, A., Mohammad Saad, M., Amzah, B., Masarudin, M. F., Jack, A., Misman, S. N., & Ramachandran, K. (2015). Buku Poket Perosak, Penyakit dan Rumpai Padi di Malaysia. Kuala Lumpur: Institut Penyelidikan dan Kemajuan Pertanian Malaysia (MARDI).

Mansor, M., Karim, S. M. R., & Hamid, Z. A. A. (2012). Seasonal variation in weed seedbank in the rice field soils of Muda area, Peninsular Malaysia. World Applied Sciences Journal, 20(8), 1086-1091.

Rabbani, N., Bajwa, R., & Javaid, A. (2011). Interference of five problematic weed species with rice growth and yield. African Journal of Biotechnology, 10(10), 1854-1862.

Rahman, M., Juraimi, A. S., Suria, A. S. M. J., Man, A. B., & Anwar, P. (2012). Response of weed flora to different herbicides in aerobic rice system. Scientific Research and Essays, 7(1), 12-23.