Rice grain classification using multi-class support vector machine (SVM)

Shafaf Ibrahim, Nurul Amirah Zulkifli, Nurbaity Sabri, Anis Amilah Shari, Mohd Rahmat Mohd Noordin
Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA Cawangan Melaka (Kampus Jasin), 77300 Merlimau, Melaka, Malaysia

Article Info

Article history:

Received May 10, 2019
Revised Jun 5, 2019
Accepted Jul 20, 2019

Keywords:

Classification
Feature extraction
Multi-class SVM
Rice grain

ABSTRACT

Presently, the demands for rice are increasing. This will affects the need for producing and sorting rice grain in faster and exceed the normal requirement. However, the manual rice classification using naked eyes are not very accurate and only professionals are able to do it. Machine learning is found to be a suitable technique for rice classification in producing an accurate result and faster solution. Thus, a study on the classification of rice grain using an image processing technique is presented. The rice grain image went through the pre-processing process which includes the grayscale and binary conversion, and segmentation before the feature extraction process. Four attributes of shape descriptor which are area, perimeter, major axis length, and minor axis length and three attributes of color descriptor which are hue, saturation and value were extracted from each rice grain image. In another note, a Multi-class Support Vector Machine (SVM) is used to classify the three types of rice grain which are basmathi, ponni and brown rice. The performance of the proposed study is evaluated to 90 testing images which returned 92.22% of classification accuracy. The study is expected to assist the Agrotechnology industry in automatic classification of rice grain in the future.

Corresponding Author:

Shafaf Ibrahim,
Faculty of Computer and Mathematical Sciences,
Universiti Teknologi MARA Cawangan Melaka (Kampus Jasin),
77300 Merlimau, Melaka, Malaysia.
Email: shafaf2429@uitm.edu.my

1. INTRODUCTION

In Asia, more than 90% of rice is produced and consumed [1]. The dependencies on rice for the majority of the population in Asia are very high and that differentiates them from the rest of the world. In Malaysia, almost all families consume rice in their daily life. The most consumed type of rice is white rice which has many nutritious values and extremely healthy. Rice paddy will go through a milling process to produce edible rice and sell to the customers. Milling is a term that consists of step by step processes of converting paddy into rice [2]. It includes cleaning, dehusking and husk separation, paddy separation, bran removal, and grading. The milling process is essential to ensure that the rice produced for customers are free from unwanted materials such as stones and sands. There are a few different types of rice milling system where it can be a simple one or two step processes, or a multi-stage process [3].

Tremendously, the demand for rice nowadays is increasing and this will affects the need for producing and sorting rice grain to produce faster and exceed the normal requirement. Rice classification refers to a process of determining and assigning rice into its classes and grade [4]. In order for the rice supplier to determine the type and grade of the rice grain, there are a few manual techniques that are carried out such as milling system, alkaline test, deoxyribonucleic acid (DNA) technique, and to name a few.

However, these manual systems have a few disadvantages. The chemical method used will cause the sample of rice grains to become destructed, and the process consumes a lot of time [5]. Consequently, this problem will slow down rice production to the market and might not meet the high amount of customer demands.

Rice classification is essential in determining the quality and its price in the market [4]. It is a process that needs to be applied in rice production industry to ensure that the rice produced for the market meets the consumer requirement. However, there are a few consumers that still in doubt with the type of rice they bought in the market. They are not satisfied until it is proven that the rice type is the same as what is stated on the rice package. The demand for quality of grains is increasing due to some traders that cheat consumers by selling poor quality food grains [6]. It is difficult for the consumers to check the quality of the type of rice that they bought as there are no suitable techniques to assist them. The only technique that consumers can use is only by using their own naked eyes inspection. Yet, researchers in [7] stated that the naked eyes inspection of the rice quality result is inefficient. DNA fingerprinting is one of the manual techniques used to determine the grades of rice, type, and quality. To maintain the uniqueness of rice varieties and to distinguish the different grades of basmathi rice [8], it is particularly important to ensure the export quality. There are many effective techniques available for extracting DNA from rice, but for milled rice samples, extraction of DNA was only possible with commercially available kits such as the Nucleon PhytoPure DNA Extraction Kit [9]. However, due to the multiple manipulation steps, this DNA extraction method is very time-consuming, laborious and expensive [10].

In this competitive modern world, people would prefer the new technologies instead of using manual and tedious techniques as mentioned. Everything is mostly now automated because of the people are requesting the things they use and expend from the technology growth [6]. In automating the process of rice grain classification, image processing technique is used to overcome the weaknesses of the manual process. Image processing is a study of an algorithm that accepts an image as the input and returns an image as output [11]. Image processing techniques are recently widely used in various applications of computer vision [12]. It is fast, economic and consistent techniques which are used in many different industries [13]. This technique satisfies the increasing production and quality requirements due to its speed and accuracy. Based on the problems discussed, the manual classification of rice grains is observed to be less accurate and it requires more experienced workers. Thus, a fast and accurate approach to classify rice grain is highly needed. Therefore, a study on the classification of rice grain using an image processing technique is proposed. The image processing technique is very effective and dependable day by day. Feature extraction techniques of color and shape were implemented to analyze the characteristics of the rice grain. Whereas, a Multi-class Support Vector Machine (SVM) technique is used to classify the three types of rice grains which are basmathi, ponni and brown rice. The outcome of this study is expected to assist the agrotechnology industry in classifying the rice grain automatically. The rest of this paper is organized as follows: Section 2 presents our research method, including the description of rice grain images, as well as the proposed feature extraction and classification techniques. Section 3 presents our results and discussions. Finally, we present our conclusions in Section 4.

2. RESEARCH METHOD

The aim of this study is to classify the rice grain using Multi-class SVM, and to evaluate the rice grain classification performance. Figure 1 depicts the proposed process flow of the rice grain classification. The proposed process of rice grain classification begins with the insertion of rice grain image. The image will then go through image pre-processing processes which are the grayscale and binary conversion, and segmentation. The next processing of processing consists of two sub-processes which are feature extraction and classification. The feature extraction process is used to extract the attributes of shape and color descriptors. Whereas, the classification is used to classify the input image and determine which rice type the image belongs to. After the image is classified, the system will produce the classification result which is the type of rice grain.

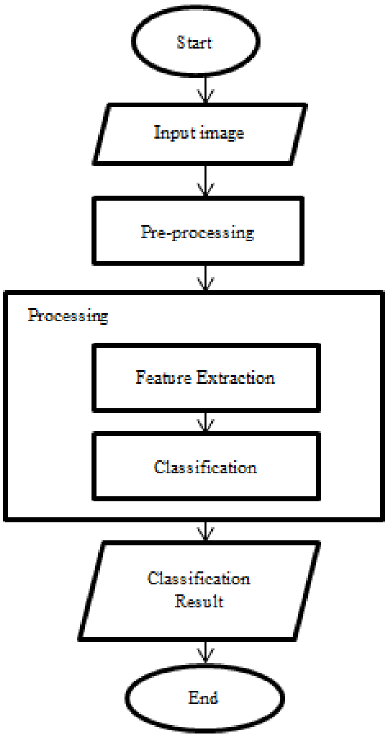


Figure 1. Proposed flowchart of rice grain classification

2.1. Input image

Ninety testing images of three types of rice grains which are basmathi, ponni and brown rice were collected. Table 1 tabulates the sample images for each type of rice grain as mentioned.

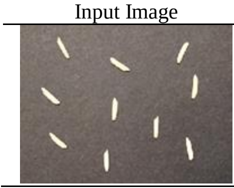
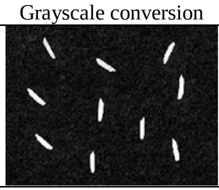
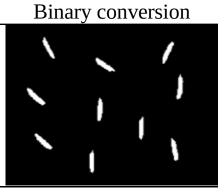
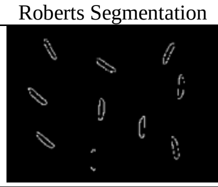
Table 1. Sample images of rice grain

Image	Type
	Basmati Rice
	Ponni Rice
	Brown Rice

2.2. Pre-processing

The pre-processing involves some processes in preparing the images for further analysis in the feature extraction process. In this study, the pre-processing includes a few processes which are greyscale and binary conversion, and image segmentation. Grayscale conversion is an approach to deal with a procedure and change the picture into dim or gray based color as the result [14]. The picture of grayscale is critical with a specific end goal to create a clean and sharp binary picture in the next procedure. In contrast, the binary image refers to a digital image which carries only two possible values for each pixel [15]. It typically has black (0) and white (1). A binary version of the processed image is created so that the toolbox function for analysis can be used. By using the *imbinarize* function, it converts the grayscale image into a binary image. The *bwareaopen* function is also used to remove background noise from the image. Image segmentation is the initial stage in image analysis [16]. An image is subdivided into its constituent parts until the objects of interest are isolated. The method that is preferred to use for this study is edge detection. Various types of edge detection can be used. For this study, the most accurate result produced is by using the Roberts edge detection technique. Table 2 depicts the sample outcomes of the grayscale conversion, binary conversion, and segmentation.

Table 2. Sample outcomes of pre-processing

Input Image	Grayscale conversion	Binary conversion	Roberts Segmentation
			

2.3. Processing

The processing part consists of two sub-processes which are feature extraction and classification.

2.3.1. Feature extraction

For the feature extraction of the rice grain, there are two types of feature that were taken into attention which are shape and color. The method that is used for shape extraction is the Regionprops function, whereas, the technique that is used for color extraction is the Hue Saturation Value (HSV). Regionprops function is used for the shape feature extraction. This function is chosen as it returns the measurements of properties specified by properties for each connected component (object) in the binary image [17]. From this function, four important shape attributes were extracted in analyzing the characteristics of the shape features of rice grain. The shape attributes are area, perimeter, major axis length, and minor axis length. Each parameter for each number of rice grain in the input image will be calculated, and the mean values are accumulated which acts as a feeder in the classification phase. In coping with the color criteria of basmathi, ponni and brown rice, the Hue, Saturation and Value (HSV) method is observed to be suitable in extracting the color attributes of the rice grain. The HSV colour space is a non-linear transformation of RGB colour space that is closed to human perception [18]. It converts the Red Green Blue (RGB) values to the appropriate hue, saturation and value (HSV) coordinates. A function of r*gb2hsv* is employed. The ranges of minimum and maximum value for each hue, saturation, and value are then were extracted.

2.3.2. Image classification

Image classification is implemented to determine the type of rice grain from the uploaded input image. A Multi-class SVM is implemented for the classification of the rice grain. The implementation of the Multi-class SVM could be done either experimentally [19] or conceptually [20]. It performs the classification by mapping the input vectors into a higher-dimensional space and building a hyper-plane that separates the data in the higher-dimensional space in an optimal way. The Multi-class SVM is chosen as it has a bigger number of classes that can be classified [21], as compared to the Support Vector Machine (SVM) which is limited to only two types of classes. In this process, the training set was used to train the Multi-class SVM model and the testing set was used to test the classification accuracy performance. The testing set is represented in (1) as follows:

$$X = \{(x_i, y_i)\}_{i=1}^{l} \text{ where } x_i \in R^n \text{ and } y_i \in \{1, 2, 3, \ldots. c\} \quad (1)$$

There are various approaches of multi-class classification problem as discussed in [22-25] such as directed acyclic graph (DAG), binary tree of SVM, one-against-all (OAA) and one-against-one (OAO). In this study, the OAO technique for multi-class classification is chosen.

2.4. Testing

The performance of the rice grain classification is evaluated using a truth table. It is performed by comparing the Multi-class SVM classification result with the actual type of rice grain. Based on the truth table obtained, the classification accuracy for each type of rice grain is calculated using (2):

$$\% \text{ of Accuracy } = \frac{No.of\ TRUE\ Accuracy\ Result}{Total\ No.of\ Testing\ Images}\ x\ 100\% \tag{2}$$

3. RESULTS AND DISCUSSION

Ninety testing images are tested for each type of rice grain. Table 3 displays some of the results plotted by the truth table from each type of rice grain. The performance of the rice grain classification is demonstrated in Table 4. From the calculation of accuracy, it is observed that the study produced a good performance with 90% of accuracy for the basmathi, and a high percentage of accuracy of 93.33% for both ponni and brown rice. The overall mean percentage of accuracy is found to produce a very good percentage of accuracy which is 92.22%.

Table 3. Rice Grain Classification

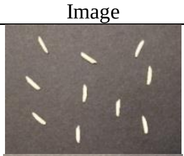
No.	Image	Rice Grain Type	Multi-class SVM classification	Accuracy Result
1		Basmathi	Basmathi	TRUE
2		Ponni	Ponni	TRUE
3		Brown	Brown	TRUE

Table 4. Accuracy Result

Rice Grain	No. of TRUE Classification	% of Accuracy
Basmathi	27	90
Ponni	28	93.33
Brown	28	93.33
MEAN		92.22

4. CONCLUSION

This paper proposed a study of rice grain classification using image processing technique. Feature extraction techniques of color and shape were implemented to analyze the characteristics of the rice grain. In another note, a Multi-class Support Vector Machine (SVM) is used to classify the three types of rice grain which are the basmathi, ponni and brown. The application to a variety of testing images has been successful. The performance of the rice grain classification is evaluated using a truth table. The overall mean percentage of accuracy demonstrated a very good percentage of accuracy which is 92.22%. It can therefore be concluded that the proposed application of image processing techniques for rice grain classification is found to be successful. Yet, implementation and incorporation of the current feature extraction and classification techniques are recommended.

ACKNOWLEDGEMENTS

The research was supported by Ministry of Education Malaysia (MoE), and Universiti Teknologi MARA through the Fundamental Research Grant Scheme **(**FRGS) (600-IRMI/FRGS 5/3 (215/2019)).

REFERENCES

[1] J. Maclean *et al.*,"Rice Almanac", Fourth Edition, Global Rice Science Partnership 2013, International Rice Research Institute. 283 P.
[2] N. Tangpinijkul, "Rice Milling System", December, pp. 1–27, 2010.
[3] P. Dhankhar, " Rice Milling", *IOSR Journal of Engineering (IOSRJEN)* , Vol. 04, Issue 05, 2014, pp. 34-42.
[4] N. Chepa *et al.*,"Determinants for Grading Malaysian Rice", *Proceedings of The International Conference on Applied Science and Technology 2016 (ICAST'16),* Vol 1761, Issue 1, 2016.
[5] J. S. Aulakh and V. K. Banga, " Grading Of Rice Grains by Image Processing", *International Journal Of Engineering Research & Technology (IJERT),* Vol. 1, Issue 4, 2012.
[6] V. Patil and V. S. Malemath, "Quality Analysis and Grading of Rice Grain Images", *International Journal Of Innovative Research In Computer And Communication Engineering*, 3(6), pp. 5679-5687, 2015.
[7] W. P. N. W. M. Tahir et al, " Rice Grading using Image Processing", *ARPN Journal Of Engineering And Applied Science* ,Vol. 10, No 21, November, pp. 10131-10137, 2015.
[8] S. Pal et al.," DNA Isolation from Milled Rice Samples for Pcrbased Molecular Marker Analysis", *Rice Genetics Newsletter*, Vol. 18, pp. 2-5, 2016.
[9] J. Thomas and V. J. Dominic, " Isolation of High Quality DNA from Pulverized Rice Grain Samples By Tri-Phase Method", *International Journal of Science and Research (IJSR),* Vol. 4, Issue 11, pp. 2013-2016, 2015.
[10] H. Liang *et al.,* "A High-Throughput DNA Extraction Method from Rice Seeds", *Biotechnology & Biotechnological Equipment*, 30(1), pp. 32-35, 2016.
[11] P. Sharma, "DSP in Image Processing", *International Journal of Advanced Research in Computer and Communication Engineering*, Vol. 4, Issue 1, January, pp. 46-49, 2015.
[12] A. Senthilrajan, " Paddy Grade And Dirt Classification using Image Processing Techniques", *2013 International Conference on Emerging Trends In Communication, Control, Signal Processing & Computing Applications (C2SPCA)*, pp. 1-3, 2013.
[13] T. Brosnan and D. Sun, " Inspection And Grading of Agricultural and Food Products by Computer Vision Systems - A Review", *Computers And Electronics In Agriculture*, Volume 36, Issues 2-3, 2010, pp. 193-213.
[14] S. A. Haney *et al.*, "An Introduction to High Content Screening: Imaging Technology, Assay Development, and Data Analysis in Biology and Drug Discovery", Wiley Publishing, ISBN: 978-0-470-62456-2, 2015.
[15] A. S. R. Sinaga, "Binary Image Object Identification Using Fuzzy Logic and Labeling Breadh-First", *Journal of Telematics and Informatics*, Vol. 6, No. 4, 2018.
[16] Kaur *et al*., "Various Image Segmentation Techniques: A Review", *International Journal of Computer Science and Mobile Computing*, Vol.3 Issue.5, pp. 809-814, 2014.
[17] C. Shanmugam and E. C. Sekaran, "IRT image segmentation and enhancement using FCM-MALO approach", *Infrared Physics & Technology*, Volume 97, pp. 187-196, 2019.
[18] S. Sural *et al*., "Segmentation and histogram generation using the HSV color space for image retrieval", *IEEE International Conference on Image Processing*, Vol. 2, 2002.
[19] R. Rifkin and A. Klautau, "In defense of one-vs-all classification", *Journal of Machine Learning Research*, Vol. 5, pp. 101–141, 2004.
[20] Y. Guermeur, "A generic model of multi-class support vector machine", *International Journal of Intelligent Information and Database Systems (IJIIDS),* Vol 6, pp. 555-577, 2012.
[21] S. B. Jadhav *et al.*, "Soybean leaf disease detection and severity measurement using multiclass SVM and KNN classifier", *International Journal of Electrical and Computer Engineering (IJECE),* Vol 9, No 5, October 2019 (Part II).
[22] C. W. Hsu and C. J. Lin, "A comparison of methods for multi-class support vector machines", *IEEE Trans. on Neural Networks*, Vol. 13, No. 2, 2002.
[23] B. Fei and J. Liu, "Binary tree of SVM: a new fast multi-class training and classification algorithm*", IEEE Transactions on Neural Networks*, vol. 17, no. 3, 2006.
[24] S. Cheong *et al.*, "Support vector machines with binary tree architecture for multi-class classification," *Neural Info. Process. Lett*., vol. 2, no. 3, 2004.
[25] D. Kasiraja and A. S. Vijendran, "Adaptive Data Structure Based Oversampling Algorithm for Ordinal Classification", *Indonesian Journal of Electrical Engineering and Computer Science,* Vol 12, No 3, December 2018.