Classification of white rice grain quality using ANN: a review

Anis Sufiya Hamzah¹, Azlinah Mohamed²

¹Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA (UiTM) Shah Alam, Malaysia
²Advanced Analytics Engineering Centre, Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA (UiTM) Shah Alam, Malaysia

ABSTRACT

Exploring the new method of using technology for classifying rice grain quality is pertinent for rice producers in order to provide quality grains and protect consumers from any contamination exist. This is even more important when in today’s market we can see that rice with low quality is traded without stringent quality control which at the end will affect consumer’s health. This paper will review classification methods in determining quality white rice grain. Although there are many researchers developing new process to do rice classification by using different technique, there are still more advanced technique that can be used to do classification. This paper will focus on classifying rice grain quality using artificial neural network (ANN) approach with the help of image processing to identify the impurities contained in the rice grains. The findings show ANN using BPNN has the highest accuracy of 96%, it is also noted that other methods provide equally better performance. This review indicate hybrid method in ANN should be explored next for future work.

Keywords:
Artificial neural network
Rice classification
Rice quality

Corresponding Author:
Anis Sufiya Hamzah
Faculty of Computer and Mathematical Science
Universiti Teknologi MARA (UiTM)
40450 Shah Alam, Selangor Darul Ehsan, Malaysia
Email: sufiya142@gmail.com

1. INTRODUCTION

Malaysia is a country based on agriculture that has many of its citizens waged as farmers for their means of living [1]. This country’s economic development also resulted from agriculture industry by providing employment in rural area. Rice is one of the most popular food that human consumes from around the world. Estimation for an adult to consume rice daily is an average of two and half plate [2]. Malaysia is one of the countries in South East Asia that produced rice along with Cambodia, Thailand and Laos. Malaysia still depends on rice imports as the country’s grain consumption is more than the self-sufficiency level (SSL) [2]. It further emphasized that this sector has to produce solution with the rise in population and demands, decreasing numbers of farmers, pesticides attacks and lack of harvesting process. There are variety of rice that can be classified according to the length, color and texture. Rice can be classified into three variety of length which are long-grain, medium-grain and short-grain (global rice science partnership).

Malaysia have valued at US$377.4 million (RM1.59 billion) or 2% of global imports which made Malaysia the world’s 14 biggest rice importer in 2016. Malaysia has given more interest in improving this number over the years. Therefore, the quality of rice needs to be monitored and maintained as it is imported from our country as consumer and producer. Good rice quality can affect the health of the consumer. In order to determine the quality of rice, many researchers came up with many solutions for classifying the rice quality. Amongst all ANN techniques, which is the best classifier that can be used for rice quality classification in...
terms of its performance, speed of training and accuracy. The objectives of this paper are to determine the quality of rice according to the rice morphological features and to find the most suitable ANN technique to be used as a classifier tool.

This paper presents a review on classification of rice grain quality using ANN technique. In this research, the goal of this paper is to reviews on the ANN techniques used as classifier for rice quality classification and discusses about the rice quality, rice morphological features as in Section 2, classifier technique and each ANN techniques performance in Section 3. The results will be discussed in Section 4 and the conclusion is written in Section 5 below.

2. RELATED WORKS

This section will introduce rice quality and also works done on classification with respect to ANN techniques in handling image processing.

2.1. What is quality?

Quality, a word that is mostly used in post-harvest studies but uncommonly touched [3]. Author also stated that quality can be mentioned as a string of values or criteria chosen on the basis of accuracy and precision of measurement. A consumer focus defines quality as consumer satisfaction, less palpable and a quantifiable concept [3]. A consumer preference on quality has its own limitations so the company needs to define their own standard of quality.

2.2. Rice quality classification

Even though Malaysia have variety of rice type, quality of the rice is most important in today’s market. Low quality rice is discolored, chalky, broken, immature, red and have been damaged. Next, the length of the rice also can be used to classify the rice quality [4]. The rice grains need to have quality requirement because consumer have to be protected from sub-standard products [5]. According to the authors, rice grains contain several impurities which are stones, straws, weed seeds, chaff, broken grains, sand and etc. This low quality rice (contains impurities) is being sold to the consumers. This somehow can affect the consumer’s health and it is still in need of improvement in grading the rice quality. However, rice type and quality classification is normally classified by experts. Classification process itself is time consuming and the volume of rice to be classified is normally in large quantity. As the technology grows, industries and people are adaptive to new technologies rather than using old techniques. According to [6], inspection of rice quality by humans is subjective and inefficient. Expert’s emotion during inspection also affects the decision making capabilities [7]. Figure 1 from [8] shows the impurities that can be contained in a bulk of rice.

Figure 1 shows a sample of impurities contained in rice grains which consists of immature and broken grains and stones. Different grades of rice may vary from the percentages of impurities found in rice grains. Usually highest grade of rice (per 1000 grams) contains less than 5 percent broken, 0.1 percent milled, 0.5 percent damaged, 0.5 percent discolored, 4 percent chalky, 0.2 percent immature, 1 percent red kernel and 10 percent paddy.

In Malaysia, Act 54, rice (grade and price control) 1992 amendment 2008 are referred to determine the rice quality and grading. In [9] classifies the objects found in Table 1 as an impurity in determining the rice quality. There are certain criteria that need to be monitored for grading the quality and identifying the superiority of rice grain. In [9] stated that the criteria of determining of the quality are composition, milling quality, damages and rice moisture. Besides, there are other criteria listed in [9] such as whitesness, brewer, translucency, milling degree, colour, foreign object, length, size, shape, chalky and width of rice.
### 2.3. Morphological rice features extractions

According to [8, 10-12], basic information of the grains such as its size and shape are used in extracting the following morphological features. Other additional features were extracted are based on the images of rice grains. These features are area which is the sum of region pixels within the seed edge, major axis length that is the longest distance of the grain from one end to the other, minor axis length is the perpendicular measurement of the distance between end points of the longest line and feature ratio is the percentage between major to minor axis length.

### 2.4. How to extract rice details for classification process?

The details in the rice need to be extracted for the purpose of results accuracy. There are several methods and technique can be used for the rice details extraction process.

#### 2.4.1. Image segmentation

Image segmentation is to extract image information. With image segmentation, important details of object can be determined and valued further within an image. As further emphasized by [13], segmentation process will lead to more informative understanding of the image.

#### 2.4.2. Image processing

The next important step to extract information from an image is to suppress or handle unwanted distortion or noise in images for further processing [13]. Hence, [13-18] used enhancement in image processing to remove noise in their image data. This enhancement intensifies the quality of images which reduce noise and blur, adjust the contrast and enlightening the image details. Below are reported steps in image processing that can be used in order to get a proper image data for the project.

a. **Gray-level co-occurrence matrix (GLCM)**
   
   GLCM is a statistical technique [5] to extract relevant features by cropping the raw size image without any change in the sample dimensions [19]. For intensity analysis, the image will convert to a grey-scale image.

b. **Color feature extraction**
   
   For color feature extraction, mean, variance, standard deviation, and range for every images will be determined [20]. In [14] also use color feature extraction to do image processing of rice image. Figure 2 shows colors that have been extracted from rice sample and can help in recognition [8].

![Figure 2 Sample of object found in rice that have been applied color feature extraction](image)

---

**Table 1. Example of defects found in grading rice quality**

| Type of defect  | Visual | Type of defect  | Visual |
|-----------------|--------|-----------------|--------|
| Big broken      | Damage | Small broken    | Immature |

---
c. Binarization

This is a process to separate background and object in which the image will be reduced to grayscale values of 0 and 1. Binarization values depending on the size of the image which starts with zero.

According to [13] and [14], to improve the classifiers performance, feature selection, is an important step that should be done before the classification process. Feature selection helps in enhancing the generalization capability.

3. CLASSIFICATION TECHNIQUES

An artificial neural network (ANN) can be defined as a model that imitating on how the human brains work [20]. Neurons in ANN are the processors that are connected to each other, that is logically the same as in the human brains. [7] said that ANNs are able to regenerate the dynamic interaction of multiple factors simultaneously which allow the research on complexity. Nowadays, ANN has step forwardly in resembling human brain. ANN has ability to learn based on the experience to improve the performance. Recognizing characters of hand-written, words identification during speech and ability to recognize moving objects are some performance of ANN. ANN also able to observe patterns in which human experts fail to do [20]. Weighting links are the connector of neurons to permit signals from neurons to neurons [20]. Weights are the base for long-term memory in ANNs. ANN is able to learn through weights adjustments. ANN also helps in processing blur images [21]. According to [22] artificial neural networks have many advantages compared to fuzzy classifiers and statistical classifiers. According to research done by [23], shows that ANN classifiers can recognize the grain based on its size and shape with an average accuracy of 98.76% and 96.67%, respectively, Next, [24] used ANN as a classifier and the percentage accuracy gives more than 96% and according to [25], the performance can reach up to 99% using ANN as classifier.

3.1. Back-propagation neural network

One of the common choices for classification in determining agricultural products is Back-propagation neural network (BPNN) [22]. BPNN is able to classify accurately poor images of varieties rice grain [16]. The researchers also had done some comparison between BPNN, SVM, KNN and Naive Bayes classifier on all data sets. The results obtained an average of 96% of accuracy by using BPNN for all features in each dataset. According to [26], based on the feature extracted, BPNN is chosen as the most suitable technique because it contains neural network (NN) layers. The NN is built with layers that give abstraction to the image. Layers are linked by weights between the range of -1 and 1 and it is randomly initialized. Training sample is used for the connection by linking and adjusting weights to meet threshold value to minimize the error value of the output with respect to the actual value [27].

3.2. Support vector machine

Another technique for classifier is support vector machine (SVM) which called supervised machine learning algorithm. According to [15] they tried using SVM to do classification but the accuracy will be increased if it has more features extracted. The accuracy results obtained from the researcher using SVM is 86%. In [4] defined that SVM is linear classifier in classifying data by learning on decision planes. The authors also mentioned that SVM shows lower percentage of classification compared to BPNN. Moreover, SVM can be a great tool, but with great power comes great responsibility [28]. There are some concerns that should be notified: First, it is very important to tune the parameters. For example, wrong choice of tuning parameters in regression can lead to much worse MSE than linear regression. Training speed for SVM is faster but the runtime complexity it is quite high for nonlinear SVM classifier [29-30]. In [30] stated that SVM shows the highest performance for rice feature selection compared to BPNN and KNN as shown in Tables 2-4.

| Table 2. Result using KNN classification [30] | Table 3. Result using SVM classification [30] |
| --- | --- |
| **precision** | **recall** | **F measure** | **precision** | **recall** | **F measure** |
| Fajr | 88.33 | 84.33 | 86.28 | Fajr | 98.33 | 98.66 | 98.49 |
| Neda | 93.33 | 95.00 | 94.16 | Neda | 99.66 | 99.33 | 99.49 |
| Hashemi | 91.00 | 96.67 | 93.75 | Hashemi | 100 | 99.33 | 99.66 |
| Mahali | 93.00 | 92.33 | 92.66 | Mahali | 98.67 | 99.00 | 97.83 |
| Gerde | 88.33 | 96.00 | 92.00 | Gerde | 99.66 | 97.66 | 98.65 |
| **Average** | 90.80 | 92.86 | 91.77 | **Average** | 99.26 | 98.80 | 98.62 |
Table 4. Result using BPNN classification [30]

|       | precision | recall  | F measure |
|-------|-----------|---------|-----------|
| Fajr  | 100       | 99.33   | 99.66     |
| Neda  | 100       | 99.30   | 99.50     |
| Hashemi | 100   | 96.6   | 98.27     |
| Mahali | 95.55  | 97.33   | 96.43     |
| Gerde | 96.66   | 95.67   | 96.16     |
| Average | 98.44 | 97.58   | 98.00     |

In the table above, using some rice morphological features, SVM shows better performance compared to other technique. In [31] stated that using SVM in classifying the rice kernels and grain, the results shown is 92.22%.

3.3. k-Nearest neighbors

Another method of classification is k-nearest neighbors (kNN). It naive solution method works well for simple recognition problems [32]. According to [33], there are several disadvantages that needs to be considered before implementing kNN. Firstly, kNN use the training data for classification unlike the other methods discussed, in which, next it is not suitable for large training data sets [33]. Lastly the major disadvantage is that kNN is sensitive to irrelevant parameters presence in the dataset.

3.4. Convolutional neural network

A convolutional neural network (CNN) is a technique that built by multi-layer neural networks which designed for visual pattern recognition through images pixel without extra preprocessing. According to [34], the results for classifying the rice disease using CNN is 91%. There are many deep learning techniques but deep CNN is the one that is mostly used for image recognition and classification [34]. The authors then said that CNN needs large amount of dataset to have better accuracy. In [34-35] used AlexNet CNN model for training the large amount of image dataset.

4. METHOD AND MATERIAL

Most of the references for this research were extracted using search engine from Science Direct, IEEE and Google Scholar. The method of search related to keywords for this review namely “rice quality”, “rice quality classification”, “neural network for classification”, “rice grading” and “rice quality classification using artificial neural network”. Each word in the title will be searched and analyzed to get the real meaning and better understanding.

The information about rice, classification and its quality can be obtained on [9] and MARDI websites. The total articles collected are 53 but only 36 articles were used as references however the rest could not be taken into consideration. It is also important to note that the papers collected for this study are current and recent to identify most relevant and recent techniques developed in rice classification thus far. Therefore, articles published from year 2005 until 2019 were chosen. All the articles and books were organized based on topic, year and domain for references. Table 5 shows summary of some papers collected for this research.

Table 5. Summary of pertinent papers

| References | Technique | Remarks |
|------------|-----------|---------|
| [4]        | BPNN      | – 93.31% using BPNN. – SVM get lower percentage compare to BPNN |
| [16]       | BPNN      | – 96% for classification – Works well with poor image of rice |
| [26]       | BPNN      | BPNN suitable for rice classification |
| [36]       | BPNN      | Overall classification accuracy achieved is 98.4%. |
| [29-30]    | SVM       | – SVM better in selected rice features – training process is faster |
| [31]       | SVM       | Able to get 92.22% for rice classification |
| [32]       | KNN       | Works on simple recognition |
| [33]       | KNN       | Not suitable for large dataset |
| [34]       | CNN       | – 91% for detection – works well with large dataset |
| [35]       | CNN       | – works well with large dataset |
| [18]       | CNN       | – works well with large dataset |
5. RESULT AND DISCUSSION

The techniques and its accuracy will be discussed in this section. Other domain in comparing the result accuracy are also studied to know which approach is the best. This section also focuses on research gap related to this topic. Figure 3 shows the comparison results of different classifier classification using chosen color features made by [4] in 2018. In [4] their research shows that by using BPNN as a method to do classification, it gives higher accuracy then SVM and k-NN for both different. It is also noted that classifier trained using PCA-based features performed better. Furthermore, BPNN classification model gives an average accuracy of 93.31% better then SVM and k-NN classifiers.

Training data set is used in NN to decrease the error in any architecture produced [36]. The architecture is then tested using testing data to compare its accuracy with training results. The input will node the value of feature extraction of rice and the output will be the type and quality of rice. From [36], the accuracy for overall classification achieved about 98.4%. In addition, [7] cited that by using BPNN technique, it can increase the accuracy and give better performances. In [15] used BPNN to compare with SVM for the classification process and the accuracy for BPNN is 96% while SVM needs more features extracted to get higher accuracy value. Most of the researchers used BPNN as classifier since it shows an increase performance from other classifier.

![Figure 3. Classification efficiency of 3 most selected methods [4]](image)

Figure 3. Classification efficiency of 3 most selected methods [4]

Figure 4 will explain the paper found for suitable classifier method related to this topic from 2014 to 2019. Figure 4 shows a graph for every 3 years on the total of related research found on the search engine. Research on classification of rice using Neural Networks are the most papers found with high accuracy compared to the other techniques. Furthermore, extra papers and research need to be done to support this proposed system in order to get better results in the future.

![Figure 4. Comparison of paper published for different classifier method](image)

Figure 4. Comparison of paper published for different classifier method

Figure 5 shows the comparison of total articles for 4 classifier methods (KNN, BPNN, SVM and others). There are 10 papers about classification of rice using neural network, 5 papers on SVM for classification, 2 papers on KNN and 2 papers for image processing and CNN that are related to this research.
Figure 5. Total papers found per 3 years

Figure 6 shows the latest method used for rice classification for the past 5 years (2015 to 2019). The latest method found to do rice classification is CNN and BPNN. In the past 5 years, they are more than 20 papers related to “rice classification” found. More than 15 researches found using BPNN as classification, 3 papers used CNN, 5 papers on SVM and less than 3 papers using KNN. The latest one published is on 2019 which use BPNN as a classifier. There is one paper published in 2019 using CNN as classifier but it was on rice diseases not rice classification.

Figure 6. Latest research on rice classification

6. CONCLUSION

The goal for this paper is to review the ANN techniques used in classification of rice quality. Quality is important in measuring the grading of rice. It is important to the Malaysian economy as Malaysia is one of the countries that imports rice from other country. Therefore, it is a need to monitor the rice quality and find the suitable technique to do the classification of rice quality. The most classification technique used based on the related works is Artificial Neural Network. Results accuracy of the past researches are mostly above 90% by using ANN technique. The accuracy result of using BPNN technique as classifier is 96% which is the best performance compared to other ANN classifier. Different researchers used different feature to measure the quality. Image processing will be done in order to get an accurate result. To help in improving performances of classifier, SVM shows the highest accuracy with less time taken to do classification. For large dataset, CNN have been proved as the best classifier and KNN is not suitable in working with large dataset. There are different features used by researchers to extract as a parameter for classifying the quality but in Malaysia, the project needs to follow the quality standard provided by BERNAS. Common technique of Artificial Neural Networks applied was BPNN to classify the rice quality. The latest method to do classification is BPNN with more than 15 papers related to rice classification are found.

ACKNOWLEDGEMENT

We would like to express our thanks to the Faculty of Computer and Mathematical Sciences, UiTM for support rendered thus far, Institute of Graduate Study, UiTM for funding and to anonymous reviewers for their useful suggestions.

REFERENCES

[1] A. Hanis, S. Jinap, M S. Nasir, R. Alias and M. S. A. Karim, “Consumers’ demand and willingness to pay for rice attributes in Malaysia,” International Food Research Journal, 19, pp. 363-369, 2012.
Classification of rice quality using ANN: a review (Anis Sufiya Hamzah)
[31] S. Ibrahim, N. A. Zulkifli, N. Sabri, A. A. Shari and M. R. M. Noordin, “Rice Grain Classification Using Multi-Class Support vector machine (SVM),” IAES International Journal of Artificial Intelligence (IJ-AI), vol. 8, issue 3, pp. 215-220, 2019.

[32] Y. Qu, C. Shang, N. M. Parthaláin, W. Wu and Q. Shen, “Multi-functional nearest-neighbour classification,” Soft Computing, vol. 22, issue 8, pp. 2717-2730, 2017.

[33] K. Amanpreet and B. Vijay, “Rice Plant Disease Detection Based on Clustering and Binarization,” Journal on Intelligent Decision Technologies, vol. 11, issue 3, pp. 357-373, 2018.

[34] R. Rommel and P. Daechul, “A Multiclass Deep Convolutional Neural Network Classifier for Detection of Common Rice Plant Anomalies,” International Journal of Advanced Computer Science and Applications (IJACSA), vol. 9, issue 1, pp. 67-70, 2018.

[35] K. S. Vimal, K. P. Monoj, M. Sonajharia and P. T Mahesh, “Rice Plant Disease Classification Using Transfer Learning of Deep Convolution Neural Network,” International Archives of the Photogrammetry, Remote Sensing and Spatial Information Science (ISPRS), vol. 3, issue 6, pp. 631-635, 2019.

[36] R. Mousavi, T. F. Akhlaghian and K. Mollazade K, “Application of imperialist competitive algorithm for feature selection: a case study on bulk rice classification,” International Journal of Computer Applications, vol. 40, issue 16, pp. 41-48, 2012.

BIOGRAPHIES OF AUTHORS

Anis Sufiya Hamzah completed her bachelors of degree at Universiti of Teknologi MARA (UiTM) in Information Technology (Hons.) Intelligent System Engineering in 2019. Currently pursuing her master’s in Information System in Intelligent Systems at the Center of Information System Studies, Faculty of Computer and Mathematical Sciences, Universiti Teknologi Mara (UiTM) Shah Alam, Selangor, Malaysia.

Azlinah Mohamed is a Professor in the Information Systems Center of Studies at the Universiti of Teknologi MARA of Malaysia. She holds a M.Sc. (Artificial Intelligence) from Bristol University and PhD (Decision Support Systems) from National University of Malaysia. Prior to this she was a tutor in University of Bristol and a Research Fellow in National University of Malaysia. Her current research interests are in the areas of Big Data, Soft Computing, Artificial Intelligence, and Web-based Decision Support Systems using intelligent agents in electronic government applications. She has published well over 140 peer-refereed journals, conference publications and book chapters internationally and locally. Besides that, she has also contributed as an examiner and reviewer to many conferences, journals and universities academic activities. Next, she had also held administration post pertaining to academic development at the university level as Head of Academic Development, Special Officer on Academic Affairs and Development to the Vice Chancellor and Dean of the Fakulti Sains Komputer dan Matematik for 7 years.