Rice Quality Detection Based on Digital Image Using Classification Method

1st Sellya Meizenty
Master of Applied Computer Engineering
Politeknik Caltex Riau
Pekanbaru, Riau, Indonesia
Sellya20s2tk@mahasiswa.pcr.ac.id

2nd Dadang Syarif Sihabudin Sahid
Master of Applied Computer Engineering
Politeknik Caltex Riau
Pekanbaru, Riau, Indonesia
dadang@pcr.ac.id

3rd Juni Nurma Sari
Master of Applied Computer Engineering
Politeknik Caltex Riau
Pekanbaru, Riau, Indonesia
juni@pcr.ac.id

***Abstract*—Rice is one of the staples that is included in the consistent list of staple food commodities (Bapok), currently some irresponsible people make the rice more durable, fragrant and whiter. Many assume that the rice is clean, odorless, and has a high price is rice with good quality and vice versa. From the existing problems the author wants to help the community to better determine good quality rice and good for consumption.**

This research will create a system that can recognize the type of rice based on the image of the rice. Rice data that has been collected will be sampled and trained using the K-Nearest Neighbors (k-NN) method where this method is used for the classification of the shortest distance calculation which will produce a class in the form of rice data classes, while to obtain parameter values from the rice image using the extraction feature. RGB color average (Red, Green, and Blue) and to get results with a good level of accuracy will use K-Fold Validation.

Keywords—Rice, RGB, K-Nearest Neighbors (k-NN), K-Fold Validation.

I. INTRODUCTION

Rice is one of the staple ingredients included in the consistent list of staple food commodities (Bapok). Staples play an important role in economic, social, and even political aspects. [15] Rice contains a lot of carbohydrates in the human body. Rice is a food that comes from the rice plant that has been separated from the skin. According to [11] 75% of the daily caloric intake of people in Asian countries comes from rice. More than 59% of the world's population depends on rice as the main source of calories

The current problem is that some irresponsible people make the rice more durable, fragrant and also whiter. So the researchers want to help the community to be more proactive about the problems that occur, namely finding the fact that a number of traders and rice entrepreneurs have found the fact that there is a bleaching agent attached to rice. so that it is known that there are fraudulent practices of a number of traders in an effort to increase prices. [16]

Many assume that rice that is clean, odorless, and has a high price is rice of good quality and vice versa, in previous studies it was assumed that rice that looks very clean, expensive and does not smell is rice of good quality.[9]

From the existing problems, the author wants to help the community to better determine good quality and good rice for consumption. This study will create a system that can recognize the type of rice based on the image of the rice, then the data that has been collected will be sampled and trained using the K-Nearest Neighbors (k-NN) method and to get the results a good level of accuracy will use k-fold validation. ice is one of the staple ingredients included in the consistent list of staple food commodities (Bapok). Staples play an important role in economic, social, and even political aspects. [15]

Rice contains a lot of carbohydrates in the human body. Rice is a food that comes from the rice plant that has been separated from the skin. According to [11] 75% of the daily caloric intake of people in Asian countries comes from rice. More than 59% of the world's population depends on rice as the main source of calories

The current problem is that some irresponsible people make the rice more durable, fragrant and also whiter. So the researchers want to help the community to be more proactive about the problems that occur, namely finding the fact that a number of traders and rice entrepreneurs have found the fact that there is a bleaching agent attached to rice. so that it is known that there are fraudulent practices of a number of traders in an effort to increase prices. [16] Many assume that rice that is clean, odorless, and has a high price is rice of good quality and vice versa, in previous studies it was assumed that rice that looks very clean, expensive and does not smell is rice of good quality.[9]

From the existing problems, the author wants to help the community to better determine good quality and good rice for consumption. This study will create a system that can recognize the type of rice based on the image of the rice, then the data that has been collected will be sampled and trained using the K-Nearest Neighbors (k-NN) method and to get the results a good level of accuracy will use k-fold validation.

II. LITERATURE REVIEW

Rice contains a lot of carbohydrates in the human body. Rice is a food that comes from the rice plant that has been separated from the skin. According to Marjuki, 75% of the daily caloric intake of people in these Asian countries comes from rice. More than 59% of the world's population depends on rice as the main source of calories [11].

The current problem is that some irresponsible people make the rice more durable, fragrant and also whiter. So the researchers want to help the community to be more proactive about the problems that occur, namely finding the fact that a number of traders and rice entrepreneurs have found the fact that there is a bleaching agent attached to rice. so that it is known that there are fraudulent practices of a number of traders in an effort to increase prices. [16].

Some researchers want to test it visually to find out the quality of rice is good and with good quality standards as well. In another study explained that currently testing is done manually so it is prone to errors. There is research on identifying rice quality with digital images [4]

This testing process is seen from the white, clean and intact value of a rice, by analyzing the value of HSV (Hue, Saturation, Value) to analyze the value of white and clean, while to analyze the value of whole rice, the object area is used. Then they took a sample of 30 data to form a decision tree classification method with the ID3 (Interative Dichotomiser Tree) model, there are 3 classification classes, namely good, poor, bad. The results of the test using the k-fold cross validation method with k = 5 obtained an accuracy of 96.67.

Other researchers studied the characteristics of rice based on analysis of image processing and artificial neural networks [1], the quality process will be seen from head rice, broken rice, groat rice and unhusked rice. researchers use digital image processing to minimize the use of inputs and parameters using index B, roundness, area, length and saturation,

To estimate brown rice, yellow rice, green rice and other foreign rice, the index parameters R, G, B, Roudness and area can be used. These values will be trained and grouped based on the type of each rice so that researchers find differences in shape, size and color so as to make different accuracy values. The results of the training between head rice, broken rice, groat rice and unhulled rice were tested with 5 input parameters the results showed a good value of 97.14% and the validation was 96.74% then for the training results from brown rice, yellow rice and green rice were 98.55 % and the validation is 90.48%.

Subsequent research discusses the same thing, namely detecting rice quality by using image segmentation based on rice grain fractions and color distribution, according to [9]. His research has factors to determine the quality of rice such as grain, non-uniform color, odor and others. This study uses the outsu feature to determine the number of broken rice grains and color distribution to determine color uniformity, this researcher uses the k-nearst neighbor classification method which has a value of 99.87 and the test results will use k-fold validation with k = 10.

The next research is on digital image processing using the SUSAN detection and Neurofuzzy methods. According to [2] the quality component of rice in his research uses a digital image processing recognition approach where the results will be able to recognize the quality components of rice which include broken grains, groats, sosoh degrees, and yellow grains. and get a score for success. The test performs blob detection on rice images with a distance of 18 cm from the camera. The sample will be tested using two methods, namely grain length using the SUSAN method, while for texture using the Neurofuzzy method.

Susan's method will include grayscale, limiting thresholding, while the Neurofuzzy method includes image extraction with GLCM to divide sata samples, training data and testing data, the resulting outputs are (good premium, bad premium, good medium, bad medium, and economical). The results obtained from this study are that the training data got good premium results, which resulted in sensitivity 51.962 %, specificity 40.151%, and accuracy 45.345%, while for testing the results were sensitivity 48.387%, specificity 38.376%, and accuracy 42.640%.

III. Research Methods

A. *Data Collection (Data Collection)*

The data used is from Bulog in Pekanbaru, Riau. Several samples of all types of rice from good quality to poor quality will be taken.

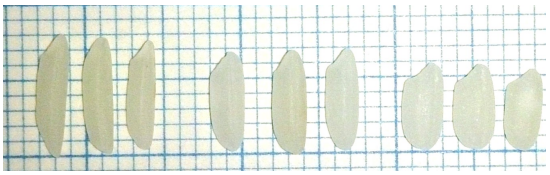

Fig. 1. Several Sample of Rice

Rice as the main food in the form of whole grains, so that shape and appearance are the first characteristics that consumers choose to choose and buy rice. The shape of the rice is a character caused by the operation of the milling process which is a combination of the type and capability of the machine, the operator's competence and the quality of the milled grain. Aspects of the shape and appearance of rice consist of length (size) and grain shape. In general, the length measurement consists of the complete length and shape based on the standard rice plant breeding process in general [7].

B. *Training Data*

K-Nearest Neighbor (KNN) belongs to the instance-based learning group. This algorithm is also a lazy learning technique. KNN is done by looking for groups of k objects in the training data that are closest (similar) to the objects in the new data or testing data.

There are many ways to measure the proximity between new data and old data (training data), including the Euclidean distance and the Manhattan distance (city block distance), the most commonly used is the Euclidean distance.

$$D(A,B)\sqrt{\sum_{k=1}^{d}(Ak-Bk)^2} \quad (1)$$

Information :

- D : proximity
- A: training data
- A: data testing
- d : the number of individual attributes between 1 to d
- f : attribute similitary function k between cases
- A and case B
- k = individual attributes between 1 to d

In finding the original nearest neighbor of the query point, K-Nearest Neighbors (k-NN) designs the nearest neighbor with the sum of the weights in each class to the query point. The winning class is determined from the smallest number of weights. The steps to calculate the K Nearest Neighbor method include:

1. Specifies the parameter (number of closest neighbors).
2. Calculate the square of the Euclidean distance (query instance) of each object to the given sample data using equation 1.
3. Then sort the objects into groups that have the smallest Euclidean distance.
4. Collect category (Nearest Neighbor Classification)
5. By using the category of Nearest Neighbor which is the majority, it can predict the value of the calculated query instance.

The image is a continuous function of the light intensity in the two-dimensional plane. The light source illuminates the object and reflects back part of the light beam [13] The image as the output of a data recording system can be optical, analog, or digital. Images that can be directly stored on a storage medium are called digital images. Digital images can be processed by a computer [10]

A digital image forms a two-dimensional field for image processing to be carried out. coordinates (x,y) represent the position of the coordinates in the Cartesian system where the horizontal axis is expressed as the x-axis, and the vertical axis is expressed as the y-axis. The light intensity function at the coordinates $(x,)$ is symbolized by $f(x,y)$ [7] Because light is a form of energy, the light intensity is between 0 and infinity, as shown in equation 4.

$$0 < f(x,y) < \infty \quad (2)$$

The intensity function (x,y) is the product of the amount of light (x,y) or illumination from the light source that illuminates the object or illumination, with the degree of the object's ability (x,y) to reflect light by the object or reflection, according to equation 5. The range of values $(x,)$ is [0,], while the range of values (x,y) is [0,1].

$$f(x,y) = i(x,y)\ .\ r(x,y) \quad (3)$$

C. Digital Image Data Acquisition

The digital image data of rice that has been collected for the testing phase is processed into textual data. The data acquired are the amount of rice, white value, net value, and rice delivery value.

1) Binary Image Segmentation

The digital image data of rice that has been collected for the testing phase is processed into textual data. The data acquired were the amount of rice, white value, net value, and whole rice value. [4].

2) Acquisition of Rice Grains

First, binary image segmentation is performed. The purpose of segmentation is to get a simple representation of an image so that it is easier to process. Segmentation is done by converting the RGB (Red, Green, Blue) rice image into a grayscale image first. Converting an RGB image to grayscale is done with the following equation: [4]

$$Grayscale = \frac{R+G+B}{3} \quad (4)$$

After being converted into grayscale image data, then it is converted to a binary image with the middle value threshold of the gray value in the image. After obtaining the binary image, segmentation is carried out by separating black pixels as background and white as objects. To facilitate object analysis for the next stage, data is collected on the location of the coordinates of each segmented object region. [4]

3) White Rice Value Acquisition

In the process of acquiring the value of white rice, analysis of the value of Hue, Saturation, and Value is carried out. The initial RGB image is converted into RGB form for the analysis process. Each value of Hue, Saturation, and Value is taken and analyzed according to the specified threshold according to the standard from the rice warehouse. The analysis process is carried out on each grain of rice according to the coordinates of the segmentation object. After all objects have been analyzed, the value is labeled as white or not white. Then calculated the percentage of objects that are white. If the number of white rice grains is less than 75% of the entire rice picture, then the rice is categorized as not white [4].

4) Acquisition of Rice Net Value

In the process of acquiring the net value of rice, the process carried out to obtain the net value is to analyze the value of Hue with a predetermined threshold according to the standards of the rice warehouse. The initial RGB image is converted into HSV format for the analysis process. The analysis process is carried out on each grain of rice according to the coordinates of the segmentation object. After all the objects have been analyzed, they are labeled as clean or unclean. Then the percentage of objects that are net worth is calculated. If the number of clean rice grains is less than 75% of the entire rice picture, then the rice is categorized as unclean. [4]

5) Rice Whole Value Acquisition

In the process of acquiring the value of the integrity of rice, it is necessary to check the area of each object first. The area calculation is done by counting the number of Mpixel in each object region. Then each object is labeled intact or incomplete value. Determination of the whole value of rice is done by comparing the object area with the standard rice area determined according to the standard from the rice warehouse. If the number of whole grains of rice is less than 75% of the whole picture of rice, then the rice is categorized as incomplete. [4]

IV. The Proposed Model

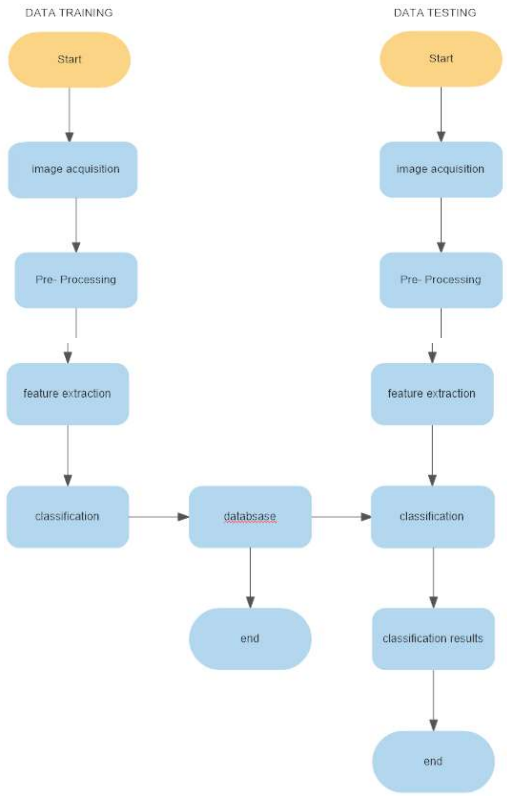

Fig. 2. Proposed Model

A. Starting from image acquisition

Then preparing the data to be tested which will then be processed. Then the parameters are taken to distinguish between other objects. Then it is entered into the system according to the existing group or characteristics. At this stage the database should be classified again and at the final stage the results of the classification will be tested with other classifications so as to get the result which is the best method..

Acknowledgment

Finally, I would like to thank all those who played a role in supporting the success of this research. This paper is far from perfect, but it is hoped that it will be useful for readers, for that the author needs criticism and suggestions from readers so that this paper is better in the future.

References

[1] Agus Supriatna Soemantri, ED (2017). Identification of Physical Quality of Rice Using Image Processing Technology and Artificial Neural Networks. Journal of Agricultural Postharvest Research (Agricultural Research and Development Agency).

[2] Amal, I. (2017). Digital Image Processing Using Susan Method. Journal of Mathematics, No.6 Vol 4.

[3] Anggraini, AP (2020, 05 18). Get to know the various types of rice and their nutritional content. Retrieved From Kompas.Com: Https://Health.Kompas.Com/Read/2020/05/18/030000368/Mengenal-Berbagai-Jenis-Beras-Dan-Kandungan-Gizinya?Page=All

[4] Arissa Aprilia Nurcahyani, RS (2015). Identification of Rice Quality With Digital Image. Scientific Journal of Informatics.

[5] Aruan, RW (2020). Comparison of the Pseudo K-Nearest Neighbor Method and the K-Nearest Neighbor Method on the Classification of Beef and Pork. Uty Journal.

[6] BV, VD (2017, 06 28). Rice. Retrieved From Indonesia Investments:Https://Www.IndonesiaInvestments.Com/Id/Bisnis/Komoditas/Beras/Item183

[7] Bbpadi, BB (2016, 01 11). Rice Quality Aspects. Retrieved From Rice Plant Science and Technology Source: Https://Bbpadi.Litbang.Pertanian.Go.Id/Index.Php/Info-Berita/Info-Teknologi/Aspek-Mutu-Beras

[8] Blp, BL (2013, 07 17-23). Harvest And Post-Harvest Rice. Agroinnovation (Sinar Tani), Pp. 6-16.

[9] Eko Supriyadi, AB (2020). Rice Quality Detection Using Image Segmentation Based On Grain Fraction And Color Distribution. Jire (Journal of Informatics & Electronic Engineering), 20-29.

[10] Coverage6. (2007, 128). Retrieved From Liputan 6: Https://Www.Liputan6.Com/News/Read/136577/Beras-Kita-Berchlorin

[11] Marjuki, FA (2008). Analysis of Factors Affecting Rice Prices in Indonesia 1981-2006.

[12] Modern, KT (2019, 11 4). Harvest and Post-Harvest Handling of Rice. Retrieved From Http://Cybex.Pertanian.Go.Id/:Http://Cybex.Pertanian.Go.Id/Mobile/Artikel/79243/Penanganan-Panen-Dan-Pasca-Panen-Padi/

[13] Nurcahyani, AA (2015). Identification of Rice Quality With Digital Image. Scientific Journal of Informatics.

[14] Pamungkas, A. (2017, August 26). Matlab Programming. Retrieved From Pemrogramanmatlab.Com/2017/07/26/Pengolahan-Citra-Digital/

[15] Prabowo, DW (2014). Clustering of Staple Food Commodities Using Analytical Hierarchy Process Method. Jurnal.Kemendag.Go.Id/.

[16] Sinabang, M. (2018, May). Analysis of Chlorine (Cl) Content in Rice and Knowledge Level of Sellers at Afternoon Market, Padang Bulan Village, Medan City in 2018.RetrievedFromHttp://Repositori.Usu.Ac.Id/Bitstream/Handle/123456789/11446/141000423.Pdf?Sequence =1&Isallowed=Y.

[17] Sofia Saida, D. (2019). Rice Quality Identification Using K-Nearest Neighbor And Support Vector Machine Method. Journal of Telecommunications, Electronics, Computing and Control.

[18] Ugm, T. (2017, 08 11). Paca Harvest Techniques. Retrieved From Post-Harvest Engineering Agricultural and Biosystem EngineeringUniversitasGadjahMada:Https://Pascapanen.Tp.Ugm.Ac.Id/2017/08/11/Teknik-Pascapanen-Padi./