Gerga Orange Quality Using Naïve Bayes Based on Feature Extraction

Fajar Raditya Nugroho*[1], **Khoirudin**[2]
Program Studi Teknik Informatika,
E-mail : fajarraditya3@gmail.com*[1], khoirudin@usm.ac.id[2]
*Corresponding author

Abstract - In an effort to increase the number of sales, the process of classifying the type of Gerga citrus fruit is very necessary. The problem that often occurs is the mixing of various types of fruit from the storage warehouse so that the quality of the fruit will be mixed and it will be difficult to determine the selling price because the quality of the fruit itself is not evenly distributed so that a sorting process is needed. There are still many sellers or growers of citrus fruits who sort the quality of the fruit manually so that it can take a very long time. Given these problems, it is necessary to classify the quality of Gerga oranges automatically with the Naïve Bayes Classifier algorithm with GLCM feature extraction and HSV color characteristics. as a method for classifying the quality of Gerga citrus fruit and as for the media used, there is digital image media. From the experiments that have been carried out the use of angles in the formation of co-occurrence matrices with the best accuracy values reaching 80% are found at angles of 0°, 45°, and 135°, while the lowest accuracy values are found at angles of 90°. It was concluded that the Gerga citrus fruit quality classification system using the Naïve Bayes method was categorized as good with an AUC value of 0.8.

Keywords - Quality, Digital Image, Classification, Naïve Bayes Classifier, GLCM

1. INTRODUCTION

In the era of information technology, digital image processing or image processing has been widely used in various fields. In the agricultural sector, disease detection, quality classification, weight determination and identification of plant species have implemented a classification process using image processing [1]–[4]. In an effort to increase the number of sales, the process of classifying the type of Gerga citrus fruit is very necessary. The problem that often occurs is the mixing of various types of fruit from the storage warehouse so that the quality of the fruit will be mixed and it will be difficult to determine the selling price because the quality of the fruit itself is not evenly distributed so that a sorting process is needed. Some significant parameters in the sorting process are size, shape, color and texture. In general, the planting and care of citrus fruits is not paid enough attention and they have not or have not implemented a production system that pays attention to quality which results in uncontrolled and low fruit quality, such as fruit skin that does not look smooth, fruit skin color tends to be green and yellow unevenly, and the sweet taste of the fruit is reduced and tastes slightly bitter, which causes growth in plant sales to not be optimal [5]. There are still many citrus fruit sellers or growers who sort the quality of the fruit manually so that it can take a very long time [6]. Given these problems, it is necessary to classify the quality of Gerga oranges automatically by utilizing image processing [7].

Digital image processing [8], [9] is the process of processing pixels in a digital image to explore and obtain specific information from the image digitally for a specific purpose. Initially the image processing process was intended as an improvement from bad image quality, but with the development of the computer world in line with the increasing capacity and speed of processing on computers as well as computational sciences that continue to grow and are very diverse, it allows researchers to retrieve information from an image to help make a decision, such as classification, identification, and facial recognition. Classification is a grouping of objects on a predetermined variable. In the classification process, digital image processing is required, image data is used in the system training process so that it can classify test data into certain groups. So a good feature extraction method is needed in order to get a good texture value [3], [10]–[12]. Naive Bayes is one of the algorithms that is often used as image classification, this algorithm classifies based on promentality and the Bayesian theorem which assumes that each variable is independent and can handle both quantitative and discrete data. Naïve Bayes also does not require large amounts of training data for the classification process so this algorithm is easy to use [13]. GLCM is a feature extraction algorithm to obtain second-order statistical values by calculating the probability of a close relationship between two pixels at a certain distance (d) and angle (θ) and the extracted features include contrast, correlation, energy, and homogeneity, according to [14] classification using the Naïve Bayes algorithm and GLCM feature extraction produces effective and efficient results. Research conducted by [15] with a discussion of comparing the Naïve Bayes algorithm with KNN to identify types of apples with LBP and HSV feature extraction with the highest results reaching 97%, namely the use of the HSV algorithm with Naïve Bayes.

2. RESEARCH METHOD

2.1. Gerga Orange

Rimau Gerga Lebong oranges or commonly called Gerga oranges, are superior local oranges that originate and are cultivated in Bengkulu province, Lebong district and Rejang Lebong district. Gerga oranges are said to be superior crops because they have a wide market share from various age groups with the support and availability of fruit that can be harvested throughout the year as shown in Figure 1. Each tree can produce bauh throughout the year.

Figure 1. Sample Dataset Gerga Orange Fruit

Gerga oranges are a type of tangerine, this orange has physical characteristics including large leaf size, fruit weight 170g-350g, fruit skin that is ready to be harvested is yellowish-green or yellow-orange in color with orange flesh color with a sweet and fresh taste, when viewed from the outside Total Dissolved Solids (TPT) is between 12-16% Brix and in terms of chemical specifications, Gerga oranges contain 89.20% water, 0.92% acid, and 18.34 mg/100 g vitamin C.

2.2. Gray Level Co-occurrence Matrix (GLCM)

Feature Extraction There are 14 types of textural features that can be extracted from the GLCM method but in this study only four main features will be used including contrast, correlation, energy and homogeneity [14], [16]. Gray Level Co-occurrence Matrix or commonly abbreviated as GLCM is a feature extraction system based on text analysis with calculations using the matrix method based on the appearance between two pixels with a certain distance, intensity, and angle in determining the degree of gray in an image. Orientation is formed in four angular directions including, 0°, 45°, 90°, and 135°. While the distance between pixels is set at 1 pixel.

- Contrast

The contrast value shows how the difference is between the dark and bright parts of an image. If the contrast value is high, the image looks sharper, whereas lowering the contrast value will reduce the sharpness of an image. Visually, the contrast value is a measure of the variation between the degrees of gray in an image area with the following equation. The higher the value of the image contrast, the higher the value of the contrast feature where P(I,j) is the co-occurrence matrix of the i-th row and j-column.

$$con = \sum_k k\char`\^2 [\sum i - \sum_i P(i,j)] \qquad (1)$$

- Homogeneity

This feature is useful for determining the level of homogeneity or uniformity between gray degrees of the same type, the more uniformity of gray degrees, the higher the value of the homogeneity feature. Where P(i,j) is the i-th row and j-column in the co-occurrence matrix.

$$homogeneity = \sum i,j \ \frac{P(i,j)}{1+|i-j|} \qquad (2)$$

- Correlation

This feature provides a clue with a value if there is a linear structure in the image by measuring the linear dependence of the gray degree on the image.

$$correlation = \sum_{i,j} \frac{(i-\mu i)(j-\mu)P((i,j)}{\sigma i \sigma j} \qquad (3)$$

Where P(i,j) : elements in the i-th row and j-column of the co-occurrence matrix, µi : average value for the i-th row, µj : average value for the jth column, σi : standard deviation value for the i-th row, j : standard deviation value for column j.

- Energy

This feature shows uniformity between neighbors on the pixel in terms of gray level, the more similar the pixel values, the higher the energy feature value will be. Where P(i,j) is the i-th row and j-column elements in the co-occurrence matrix.

$$energy = \sum_{i,j=0}^{N-1} P_{(i,j)\char`\^2} \qquad (4)$$

- RGB Color Space

Basically, color images come from a color space that comes with a combination of the three main colors namely red, greed, and blue and are combined to produce a variety of different colors. The color combination model also depends on the device used and allows each device to detect and produce different color combination values. The following are the color parameters obtained by normalizing each RGB component in the image with the following equation.

$$r = \frac{R}{R+G+B} \quad g = \frac{G}{R+G+B} \quad b = \frac{B}{R+G+B} \tag{5}$$

HSV is a color feature in digital images which, among other things, represents HUE, Saturation, and Value values. HUE is a basic color identity that differs from one color to another expressed in degrees, namely Red 0°, Green 120° and Blue 240°. For example, oranges are generally yellow in color, different from mangoes which are green, this color identity difference is called HUE. Saturation is the level of density of a color, the higher the level of density of a color, the closer it is to the base color, the lower the level of density, the color will be more faded and gray in color [15]. Value is the brightness level of a color, the brighter a color, the higher the V value. A bright color has more white elements while a dark color has more black elements.

$$H = \begin{cases} 0°, jika\ V = 0 \\ 60°.\left(\frac{G-B}{vm} mod\ 6\right), jika\ V = R \\ 60°.\left(2 + \frac{B-R}{vm}\right), jika\ V = G \\ 60°.\left(4 + \frac{R-G}{vm}\right), jika\ V = B \end{cases} \tag{6}$$

$$S = \begin{cases} 0, jika\ V = 0 \\ \frac{Vm}{V}, jika\ V > 0 \end{cases} \tag{7}$$

$$V = 1/3(R + G + B) \tag{8}$$

2.3. Naïve Bayes

The Naïve Bayes algorithm is one of the algorithms found in classification techniques. This algorithm was discovered by a British scientist named Thomas Bayes in 1763, Naive Bayes is a classification using probability and statistical methods, this algorithm is called Bayes' Theorem because it predicts opportunities based on variables and previous experience with new variables being tested [17]–[19]. The theorem is combined with Naïve which means that all conditional attributes are free. Classification with this algorithm will assume a class characteristic for a particular class and has nothing to do with other classes as shown I equation (9).

$$P(X/Y) = \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-(x-\mu)^{\wedge}2}{2\sigma^{\wedge}2}} \tag{9}$$

Where P(X|Y): Likelihood of Occurrence in One Class, X : Parameters, σ^2 : Parameter Standard Deviation, "e" : Exponential, and μ : Parameter Average Value.

$$P(H/X) = \frac{P(X/H)P(H)}{P(X)} \tag{10}$$

Where X : Classless X data to test, H : Assumes the ith data into one of the classes, P(H|X) : Probability of hypothesis H based on condition X, P(H) : Probability of the hypothesis H, P(X|H)

: Probability of hypothesis X based on condition H, and P(X) : Probability of hypothesis X. The flow of the Naïve Bayes algorithm is as follows:

- Reading test data by parameters
- Determine the mean and standard deviation values for each parameter
- Calculating the probability of the value appearing on each parameter
- Calculating test data for each probability class based on training data
- Determine the highest probability value

2.4. Confusion Matrix

The Confusion Matrix is one of the test systems to measure the level of accuracy of a prediction based on the actual value of an expert. The Confusion Matrix Single Decision Threshold will be used to measure how the performance of the predicted classification results is in accordance with the actual value of the quality of Gerga oranges. There are several terms in the Single Decision Threshold method, including:

- True Positive (TP) is where the condition has a real positive value on the test data and will produce a positive value prediction on the system.
- True Negative (TN) is where the condition has a negative real value on the test data and will produce a negative predicted value on the system.
- False Positive (FP) is where the condition has a negative real value in the test data but produces a positive value in the system.
- False Negative (FN) is where the condition has a positive real value on the test data but produces a negative value on the system.

For application to the system, parameters will be used, namely sensitivity, precision, specificity, and accuracy [15][16]. Sensitivity will know how many positive values but with negative predictive results. Precision is useful for getting positive test values with positive predictions. Accuracy that is useful for calculating the success rate of global classification on the system. Specificity is the predicted negative value with the true value being negative as illustrated in equation (11) until (14).

$$Sensitivity = TP/(TP+FN).100\% \qquad (11)$$

$$Precision = TP/(TP+FP).100\% \qquad (12)$$

$$Accuracy = (TP+TN)/(TP+TN+FP+FN).100\% \qquad (13)$$

$$Specificity = TN/(TN+FP).100\% \qquad (14)$$

The predicted value is the value generated from the system, while the real value is the value that corresponds to the events that occur or is referred to as the expert value. The following is the Confusion Matrix which has been described in Table 2 and for the Confusion Matrix and the formula that will be used according to the class label classification in this final project as follows in equation (15).

Table 1. Confusion Matrix

		Expert Value	
		True	*False*
Value Predict	*True*	TP	FP
	False	FN	TN

Table 2. Confusion Matrix Quality Classification of Gerga Citrus Fruits

		Expert Value		
		The Best	Good	Not Good
Value Predict	The Best	TP	FP1	FP2
	Good	FC1	TC	FC2
	Not Good	FK1	FK2	TK

Information :

- TP : System testing for the best image where expert values and test results match.
- FP1: Inappropriate system testing where the expert value of the test data is quite good with the best predictive value.
- FP2: Inappropriate system testing where the expert value of the test data is not good with the best predictive value.
- TC : System testing for images is good enough where expert values and test results match.
- FC1: Inappropriate system testing where the expert value of the test data is the best with a pretty good predictive value.
- FP2: Inappropriate system testing where the expert value of the test data is not good with the best predictive value.
- TK: Prediction according to the test image is not good and the expert value is not good.
- FK1: Inappropriate system testing where the expert value of the test data is the best with a less good predictive value.
- FK2: Inappropriate system testing where the expert value of the test data is quite good with a poor predictive value.

$$\text{"Precision ="} \ \text{"TP"} \ / \ \text{"TP + FP1 + FP2"} \ \text{"100\%"} \qquad (15)$$

$$\text{"Sensitivity Fairly Good = "} \ \text{"TC"} \ /\text{"TC + FC1 + FC2"} \ \text{"100\%"} \qquad (16)$$

$$\text{"Poor Sensitivity = "} \ \text{"TK"} \ /\text{"TK + FK1 + FK2"} \ \text{"100\%"} \qquad (17)$$

$$\text{"Accuracy ="} \ \text{"TP + TC + TK"} \ / \ \text{"Sum of Test Data"} \ \text{"100\%"} \qquad (18)$$

As a benchmark for the performance and feasibility level of the classification program that will be made using the Area Under Curve (AUC). Based on the accuracy value, it will be concluded whether the Gerga citrus quality classification system can be applied properly and can be used properly. For the value of the level of accuracy in the AUC itself between 0.0 to 1.0 the more the resulting value is close to the value of 1.0, the classification is said to be very good or very feasible and if the resulting accuracy value is less than 0.6 then it can be said that the classification method used is not good or wrong [20], [21]. In the following, the categories of AUC values have been described from the best value to the less good value in Table 4 and the equation is as follows.

Table 3. AUC Categories

AUC Value	Information
0,91 – 1,00	Paling Bagus
0,81 – 0,90	Bagus
0,71 – 0,80	Cukup Bagus
0,61 – 0,70	Kurang
<= 0,60	Salah

$$\text{"}AUC = \text{" "}1 + Sensitivity - (1 - Specificity)\text{" /"}2\text{"} \qquad (19)$$

2.5. Classification Scheme

In the training process, each training data has been labeled into one class, namely the best, pretty good, and not very good, which will be calculated for the features used to create the training data database as benchmark values in the testing process. The process of testing the classification system is carried out with test images based on benchmark values in the training data database, the system will return the test image values into one of the best, sufficient, or not good classes as shown in Figure 2.

Figure 2. Training Stages

This process is also useful to find out whether the classification system is working properly and the process of calculating the accuracy value is also carried out. The training stage is the initial stage to obtain training characteristics and training targets on training images with the final result being a training database model that will be used in the test image classification process as a reference value based on expert values with the Naïve Bayes algorithm. Calling or selecting images is the initial stage of the classification process at the training and image testing stages. The training images used are 300 images with jpg extension which have been grouped into the best, moderately good, or less good classes according to the real conditions where each class has 100 training images. The process of calling a training image is shown in Figure 3.
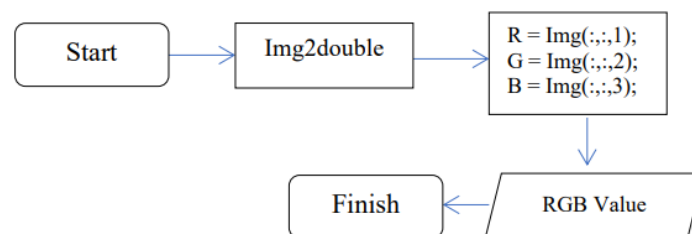

Figure 3. RGB Feature Extraction

The process begins by selecting training images that are already in a folder that has been grouped into one class, the image extension used is .jpg if the extension is not .jpg then the image will be ignored, if the extension is .jpg then the process of calculating the total images that will be used for the stage next. This preprocessing stage is carried out after calling the

image, whether it is a training image or a test image, there are two processes in preprocessing, namely reducing the size or resizing the image and converting the color image to grayscale.
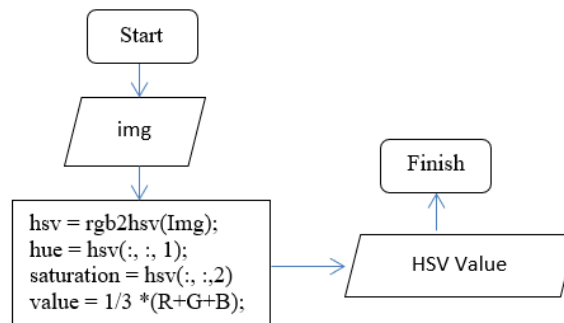

Figure 4. HSV Feature Extraction

The image resizing stage aims to speed up image processing time, the image resizing process is carried out after calling a training image or test image, the initial image size is 2048 x 1536 pixels, the image will be reduced to 256 x 256 pixels. Image conversion is carried out to convert a color image into a grayscale image to obtain the gray level of an image, in contrast to a color image which has three color components, namely red, green, and blue, whereas a grayscale image only has one component, namely gray with a scale of 0 to 255, this image is will be used in GLCM calculations to obtain texture features on the training image and test image.
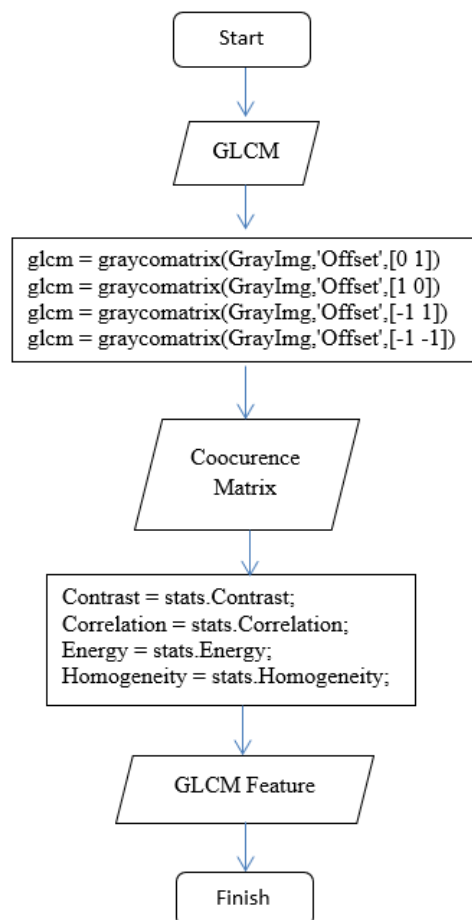

Figure 5. GLCM Scheme

Hue (H) is the basic color of an image which is described in degrees, namely 0 degrees red, 120 degrees green, and 240 degrees blue. Saturation (S) is the level of density of a color, the higher the level of density of a color, the closer it is to the base color, the lower the level of density, the color will be faded and gray in color. Value (V) is the brightness level of a color, the brighter a color, the higher the V value as shown in Figure 4. A bright color has more white elements while a dark color has more black elements. Each image that has been processed to form a co-occurrence matrix will perform GLCM feature calculations to obtain its features, namely Contrast, Correlation, Energy and Homogeneity, each image from the best class, good enough, and not good enough will get these four features. The GLCM feature extraction process is as shown in Figure 5.
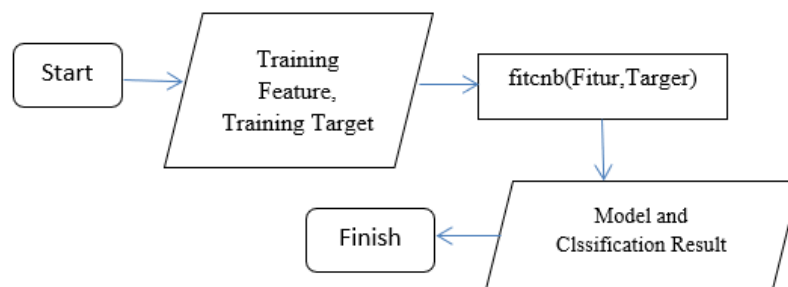
Figure 6. Naïve Bayes Proposed Scheme

Based on Figure 6, extraction of training data image features is stored in a database in .xls format, this database is carried out by the Naïve Bayes trainer process and a database of modeling results is obtained in .mat format. Structure of the training database. After the feature extraction process and training on the training data is complete, the next process is testing with the test data image. This stage is carried out as a benchmark for how effective the classification system with the algrotima and training images used is. Tests were carried out with 30 test data images and the angles used in the GLCM feature extraction were 0°, 45°, 90°, and 135° as shown in Figure 7.
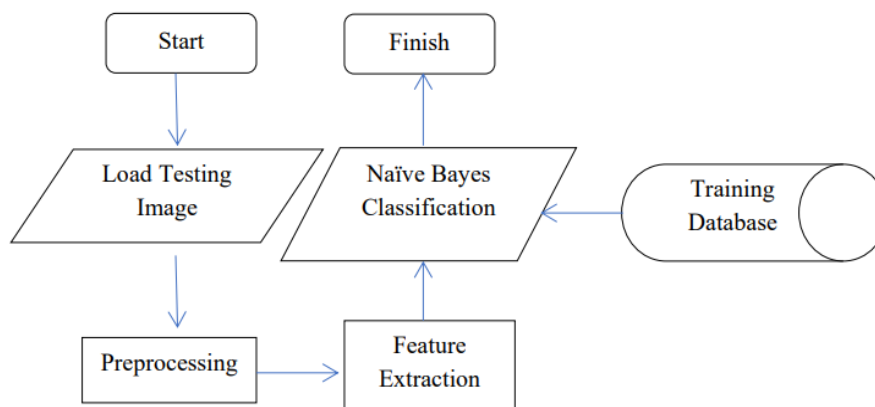
Figure 7. Testing Stages

3. RESULTS AND DISCUSSION

Graphic User Interface (GUI), which is a display that will bridge between computer programs and users, the use of a GUI is needed to make it easier for users to operate a system on a computer, namely the quality classification system of Gerga oranges in this study. The part

that will be implemented into the GUI is the testing phase, at this stage it will explain every part of the GUI that is made. The GUI display is as shown in the following Figure 8. The training database named NaiveBayes is called with the load function and the classification process is carried out with the predict function based on the test characteristics with the training targets stored in the training database, if the classification results return number 1 then the classification result is the best class, if the number is 2 then the class is good enough, and if the classification result returns number 3 then the classification result is a poor class, the classification results are displayed on the Classification Results panel in the GUI with data of type String.
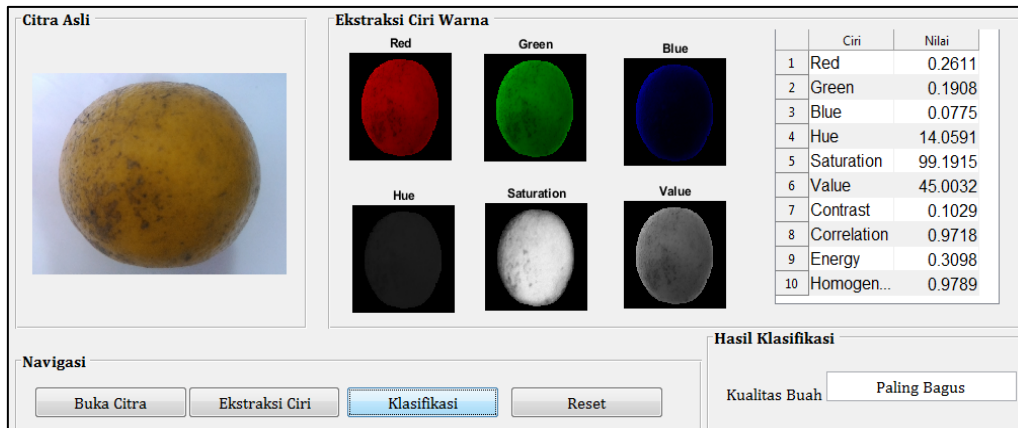

Figure 8. GUI Application Gerga Orange Quality

Based on the classification results with the Naïve Bayes algorithm, two methods will be tested, namely the Confusion Matrix and Area Under Curve (AUC). Testing with the Confusion Matrix method aims to determine how accurate the system is in determining the correct classification of test image quality, while the AUC method aims to measure the performance of the classification method used in this study, namely the classification method with the Naïve Bayes algorithm. Based on the experiments that have been carried out the use of angles in the formation of the co-occurrence matrix with the best accuracy value reaching 80% is found in the use of angles 0°, 45°, and 135°, while the lowest accuracy value is found in the use of angles of 90° with a value of 73.3333% as shown in Table 4.

Table 4. Classification accuracy using Naïve Bayes

	GLCM Degree											
	0°			45°			90°			135°		
Class	The Best	Good	Not Good	The Best	Good	Not Good	The Best	Good	Not Good	The Best	Good	Not Good
The Best	10	0	0	10	0	0	10	0	0	10	0	0
Good	3	6	1	3	6	1	4	4	2	3	4	1
Not Good	0	2	8	0	2	8	0	2	8	0	2	8
	Percentage (%)											
Accuracy	80			80			73,333			80		
Sensitivity	100			100			100			100		
Specificity is Good Enough	60			60			40			60		
Less Good Specificity	80			80			80			80		

This test is carried out to measure how effective the classification method used is. What is the AUC value in the range of 0 and 1, if the value gets closer to 1 then the method used can be said to be effective in classifying images. The AUC calculation is done with the following equation (18).

$$AUC=((1+Sensitivity)-(1-Spesivicity))/2 \qquad (18)$$

Testing this classification method will use the values obtained from the Confusion Matrix test with the results obtained as shown in Table 5.

Table 5. Testing Method using AUC

Degree	AUC Results
0°	0.85
45°	0.85
90°	0.8
135°	0.85

Based on the tests carried out there were several classification results that were not in accordance with expert scores, with the lowest accuracy value reaching 73.3333% and the AUC value reaching 0.8 actually the system was categorized as good, the most classification errors were found in pretty good classes with the highest Specificity value only reaching 60%. From the analysis of system failures in conducting classification, there are several conclusions, namely taking images with poor and inconsistent light conditions, this causes the color of the image to change, tends to be dark, reddish, or bluish, too bright light also causes the image background to be too shady as shown in Figure 8.

Figure 8. Shady Image

The condition of an image that is out of focus also makes the image look blurry, this makes the conversion of the gray level in the image not good and the feature extraction results are not optimal as shown in the following Figure 9. Poor lighting conditions cause the background of the image to be shaded, making the color feature extraction less irregular. Under these conditions, the color feature extraction results become excessive and inaccurate.
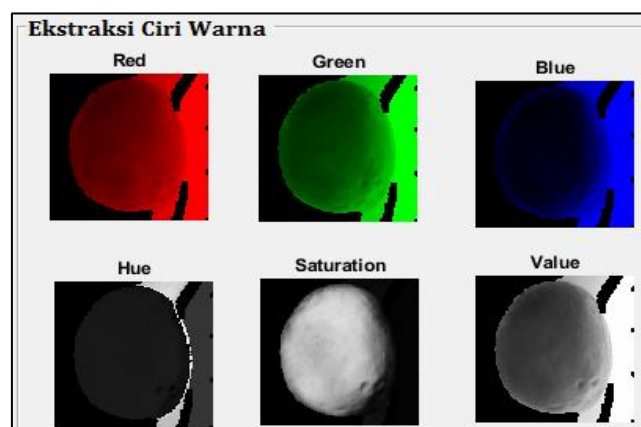

Figure 9. The condition of the test image is not good

4. CONCLUSION

Based on the research that has been done, it can be concluded that the use of the Naïve Bayes classification algorithm based on GLCM feature extraction in the Gerga citrus fruit quality classification system with image media is categorized as good and successful with the highest AUC value reaching 0.85. In this study there are still some deficiencies that need to be considered and corrected so that suggestions are needed for further research. The following are suggestions that need to be considered.

a. To take training images or test images, attention should be paid to the light conditions so that the fruit images are not too shady so that the feature extraction process can be appropriate and maximized.
b. The use of more training images and the use of lighting aids are expected to increase the accuracy value in the image classification process.
c. In the testing process using only one test image so it will be less effective if it is carried out in the sorting process at Gerga Orange plantations so it is hoped that in the future a fruit quality classification machine will be made using many test images at one time based on the method used in this study.
d. This research is expected to help and motivate researchers and farmers of Gerga oranges or other oranges to develop a system for sorting the quality of oranges in the hope of increasing the sales quality of the oranges themselves so that they can contribute to the development of sales and cultivation of oranges in Indonesia.

REFERENCES

[1] E. K. Ratnasari and A. Wikaningrum, "Pengenalan Jenis Buah pada Citra Menggunakan Pendekatan Klasifikasi Berdasarkan Fitur Warna Lab dan Tekstur Co- Occurrence," *J. Inf.*, vol. 1, no. 2, pp. 88–97, 2016.

[2] P. B. Padol and A. A. Yadav, "SVM classifier based grape leaf disease detection," *Conf. Adv. Signal Process. CASP 2016*, pp. 175–179, 2016.

[3] H. Cecotti, A. Rivera, M. Farhadloo, and M. A. Pedroza, "Grape detection with convolutional neural networks," *Expert Syst. Appl.*, vol. 159, p. 113588, 2020.

[4] Y. Ji, Q. Zhao, S. Bi, and T. Shen, "Apple Grading Method Based on Features of Color and Defect," in *Chinese Control Conference*, 2018, pp. 5364–5368.

[5] D. Nursantika and F. R. Umbara, "PENGENALAN CITRA BUAH MANGGIS MENGGUNAKAN METODE JARINGAN SYARAF TIRUAN BACKPROPAGATION," in *Seminar Nasional Telekomunikasi dan Informatika*, 2016, no. Selisik, pp. 182–184.

[6] S. Musacchi and S. Serra, "Apple fruit quality: Overview on pre-harvest factors," *Sci. Hortic. (Amsterdam).*, vol. 234, no. December, pp. 409–430, 2018.

[7] K. Komal and Sonia, "GLCM Algorithm and SVM Classification Method for Orange Fruit Quality Assessment," *Int. J. Eng. Res. Technol.*, vol. 8, no. 09, pp. 697–703, 2019.

[8] I. U. W. Mulyono *et al.*, "Parijoto Fruits Classification using K-Nearest Neighbor Based on Gray Level Co-Occurrence Matrix Texture Extraction," *J. Phys. Conf. Ser.*, vol. 1501, no. 1, 2020.

[9] T. Y. Prahudaya and A. Harjoko, "Metode Klasifikasi Mutu Jambu Biji Menggunakan Knn Berdasarkan Fitur Warna Dan Tekstur," *J. Teknosains*, vol. 6, no. 2, p. 113, 2017.

[10] P. Rianto and A. Harjoko, "Penentuan Kematangan Buah Salak Pondoh Di Pohon Berbasis Pengolahan Citra Digital," *IJCCS (Indonesian J. Comput. Cybern. Syst.*, vol. 11, no. 2, p. 143, 2017.

[11] S. F. Kusuma, R. E. Pawening, and R. Dijaya, "Otomatisasi Klasifikasi Kematangan Buah Mengkudu Berdasarkan Warna Dan Tekstur," *Regist. J. Ilm. Teknol. Sist. Inf.*, vol. 3, no. 1, pp. 17–23, 2017.

[12] S. Y. R. Riska and P. Subekti, "Klasifikasi Level Kematangan Buah Tomat Berdasarkan Fitur Warna Menggunakan Multi-Svm," *J. Ilm. Inform.*, vol. 1, no. 1, 2016.

[13] A. Ciputra, D. Rosal, I. M. Setiadi, E. H. Rachmawanto, and A. Susanto, "KLASIFIKASI TINGKAT KEMATANGAN BUAH APEL MANALAGI DENGAN ALGORITMA NAIVE BAYES DAN EKSTRAKSI FITUR CITRA DIGITAL," *J. SIMETRIS*, vol. 9, no. 1, 2018.

[14] O. R. Indriani, E. J. Kusuma, C. A. Sari, E. H. Rachmawanto, and D. R. I. M. Setiadi, "Tomatoes classification using K-NN based on GLCM and HSV color space," in *2017 International Conference on Innovative and Creative Information Technology (ICITech)*, 2017, pp. 1–6.

[15] P. N. Andono, E. H. Rachmawanto, N. S. Herman, and K. Kondo, "Orchid types classification using supervised learning algorithm based on feature and color extraction," *Bull. Electr. Eng. Informatics*, vol. 10, no. 5, pp. 2530–2538, Oct. 2021.

[16] A. Nosseir and S. E. A. Ahmed, "Automatic Classification for Fruits' Types and Identification of Rotten Ones Using k-NN and SVM," *Int. J. Online Biomed. Eng.*, vol. 15, no. 03, p. 47, Feb. 2019.

[17] F. Liantoni and H. Nugroho, "Klasifikasi Daun Herbal Menggunakan Metode Naïve Bayes Classifier Dan K- Nearest Neighbor," *Junral Simantec*, vol. 5, no. 1, pp. 9–16, 2015.

[18] N. Ani Brown Mary and D. Dharma, "A novel framework for real-time diseased coral reef image classification," *Multimed. Tools Appl.*, vol. 78, no. 9, pp. 11387–11425, May 2019.

[19] G. Saleem, M. Akhtar, N. Ahmed, and W. S. Qureshi, "Automated analysis of visual leaf shape features for plant classification," *Comput. Electron. Agric.*, vol. 157, no. November 2018, pp. 270–280, 2019.

[20] S. H. Wady, R. Z. Yousif, and H. R. Hasan, "A Novel Intelligent System for Brain Tumor Diagnosis Based on a Composite Neutrosophic-Slantlet Transform Domain for Statistical Texture Feature Extraction," *Biomed Res. Int.*, vol. 2020, pp. 1–21, Jul. 2020.

[21] J. Amin, M. Sharif, M. Raza, T. Saba, and M. A. Anjum, "Brain tumor detection using statistical and machine learning method," *Comput. Methods Programs Biomed.*, vol. 177, pp. 69–79, Aug. 2019.