

Classification of Tangerine (*Citrus Reticulata Blanco*) Quality Using Combination of GLCM, HSV, and K-NN

Friska Ayu Listya Irawan^{*1}, Nur Rokhman^{*2}

¹Master Program of Computer Science; FMIPA UGM, Yogyakarta, Indonesia

²Department of Computer Science and Electronics, FMIPA UGM, Yogyakarta, Indonesia

e-mail: ^{*1}friska.a@mail.ugm.ac.id, ²nurrokhman@ugm.ac.id

Abstrak

Kualitas hasil produksi buah merupakan hal yang sangat penting karena terkait dengan nilai penjualan. Data dari Direktorat Jenderal Hortikultura Kementerian Pertanian pada tahun 2017 menunjukkan bahwa 94,3% dari keseluruhan hasil produksi buah jeruk merupakan jenis jeruk Keprok. Dalam klasifikasi kualitas jeruk keprok proses pengamatan visual sangat dipengaruhi oleh subjektivitas sehingga pada kondisi tertentu seperti mata lelah maupun jumlah jeruk yang ingin klasifikasi sangat banyak prosesnya bisa tidak konsisten dan juga memakan waktu yang lama. Oleh karena itu, dibutuhkan suatu teknologi untuk mempercepat proses klasifikasi dan membuatnya lebih objektif. Penelitian ini menggabungkan metode Gray level Co-occurrence Matrix (GLCM) untuk fitur tekstur, Hue, Saturation, Value (HSV) untuk fitur warna dan metode klasifikasi k-Nearest Neighbor (k-NN). Data yang digunakan adalah 60 citra jeruk keprok busuk dan 60 citra jeruk keprok tidak busuk dibagi menggunakan metode 4-fold cross-validation untuk mendapatkan kombinasi data latih dan data uji terbaik. Terdapat 3 proses utama yang akan dilakukan yaitu preprocessing, ekstraksi fitur dan klasifikasi. Penelitian ini menghasilkan akurasi tertinggi 80% dari gabungan ekstraksi fitur GLCM dan HSV dengan pengujian nilai $k = 5$ pada k-NN.

Kata kunci—Klasifikasi, Jeruk Keprok, GLCM, HSV, k-NN

Abstract

The quality of fruit production is very important because it is related to the value of sales. Data from the Directorate General of Horticulture at the Ministry of Agriculture in 2017 showed that 94,3% of the total yield of citrus fruits is a type of tangerine. In the classification of the quality, the visual observation process is strongly influenced by subjectivity so that in certain conditions such as tired eyes and the number of oranges that want to classify too many the process can be inconsistent and also take a long time. Therefore, a technology is needed to accelerate the classification process and make it more objective. This study combines the Gray level Co-occurrence Matrix (GLCM) method for texture, Hue, Saturation, Value (HSV) features for color features and the k-Nearest Neighbor (k-NN) classification method. The data used were 60 images of rotten tangerines and 60 images of not rotten tangerines divided using a 4-fold cross-validation method to find the best combination of data training and data testing. 3 main processes will be carried out, namely preprocessing, feature extraction and classification. This study produced the highest accuracy of 80% from the combined of GLCM and HSV features extraction with value $k = 5$ for k-NN.

Keywords— Classification, GLCM, HSV, k-NN

1. INTRODUCTION

The quality of fruit production is very important related to the value of sales. Because one of the important roles in national economic development is horticultural commodities. Data from the Directorate General of Horticulture at the Ministry of Agriculture in 2017 showed that Tangerines had a production of 2.165.184 tons of total citrus fruit production of 2.295.310 tons. This shows that 94,3% of the total citrus fruits are tangerines [1]. Fruit quality can be seen from the size, color, and defects on the skin. Today, most quality classifications are still done manually (visual observation) by farmers and sellers. For citrus farmers, it is easy, but it is still undeniable because the manual method is very dependent on subjectivity such as dry eyes and the number of oranges that want to be sorted out is too many, the process will be inconsistent and also requires a long time.

Digital image processing is carried out on the image to get certain results according to needs. With digital image processing, the image of oranges can be processed to determine the quality of oranges including rotten or not rotten oranges. This digital image processing can be one way to solve problems that have been stated before. Feature extraction is one of the important things in digital image processing. With feature extraction, we can know the characteristics that represent each image.

Many previous studies have used digital image processing for fruit classification processes. Starting from the quality and maturity level. Some of them are the classification of the quality of tangerines using the Gray Level Co-occurrence Matrix (GLCM) method for texture features. The quality of tangerines is classified into 3 classes namely Grade A, Grade B, and Grade Super. Then use the Support Vector Machine (SVM) method to identify whether or not an image windowing area is defective. This study resulted in the best accuracy rate of 82.5% with the number of training data as many as 20, the distance value is 2 in the direction of GLCM 45° [2]. There is also a classification of the maturity level of lime-based on color features and k-Nearest Neighbor (k-NN) classification. Color features use mean RGB values, and the k values used are 1, 3, 5, 7, and 9. Search for the closest distance between training data and test data using Euclidean Distance and Cityblock Distance. This study produced the highest accuracy of 92% with Euclidean Distance and the value of $k = 3$ was the best k value [3]. And also, research using a combination of feature extraction is a classification of Tomatoes using k-NN based on GLCM and HSV color space. The result using 100 data sets, consisting of 75 training data and 25 testing data yields the highest accuracy rate of 100% with p-value on GLCM is 9 and the membership value (k) in k-NN is 3. According to the experimental results, we can conclude that the proposed method can achieve the highest accuracy [4].

From several previous studies that have described the combination of GLCM and HSV features get the highest accuracy, therefore in this study, GLCM will be used as the extraction of feature texture and HSV color space as the extraction of color features. Classification using the k-NN method for classification of the quality of tangerines rotten and not rotten. This research is expected to accelerate the classification process and make it more objective.

2. METHODS

2.1 Digital image processing

Digital image processing is a technical step in estimating the characteristics of objects in the image, measuring the characteristics related to the geometry of the object and interpreting the geometry [5]. Formation of digital images through several stages, namely image acquisition, sampling, and quantization. Image acquisition is the process of making continuous images into digital forms. Images are obtained as the initial process of image processing. The tool used for image acquisition is a device equipped with sensors that can capture images and convert light energy into digital signals (camera, scanner, etc.). The sampling process is also called the digitizing process at the coordinate x, y. digitization is the transformation of continuous images

into digital images by dividing analog (continuous) images into M columns and N lines so that they become discrete images. The greater the value of M and N, the finer the digital image is produced and this also shows the higher the resolution of the image. The quantization process is produced by digital equipment (scanners, digital cameras, etc.). Analog intensity transformation, namely the value of continuous amplitude to the area of discrete intensity is called quantization. The measured amplitude value is the value in the discrete coordinates of the sampling process [6]. In this research, as shown in Figure 1 will use the image of rotten tangerines and the image of tangerines are not rotten.

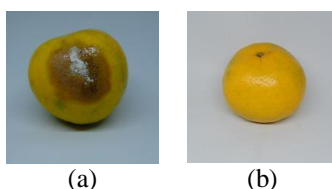


Figure 1 (a) Rotten tangerines, (b) Tangerines are not rotten

2. 2 Features Extraction

Feature extraction is one of the important things in digital image processing. With feature extraction, we can know the characteristics that represent each image. In this research will use extraction feature for texture using Grey Level Co-occurrence Matrix (GLCM) method and for feature color using Hue, Saturation, Value (HSV) method.

2. 2.1 Gray Level Co-occurrence Matrix (GLCM)

GLCM is one of the advanced techniques used for texture feature extraction was proposed by Haralick back in 1973, It shows how often a pixel value is known as the reference pixel with the intensity value i occurs in a specific relationship to a pixel value known as the neighbor pixel with the intensity value j . So, each element (i,j) of the matrix is the number of occurrences of the pair of a pixel with the value i and a pixel with value j which are at a distance d relative to each other. The spatial relationship between two neighboring pixels can be specified in many ways with different offsets and angles, four possible spatial relationships (0; 45; 90 and 135). Figure 2 illustrates the details of the process to generate four symmetrical co-occurrence matrices with one neighboring pixel ($d=1$) along four possible directions as $\{[0\ 1]$ for 0; $[-1\ 1]$ for 45; $[-1\ 0]$ for 90 and $[-1\ -1]$ for 135} [7].

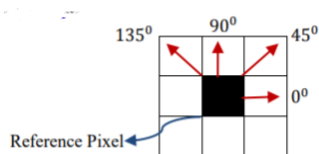


Figure 2 Co-occurrence matrix directions for extracting texture features

The co-occurrence matrix captures the texture properties but cannot be directly used as an analytical tool, for example comparing two textures, this data must be extracted again to get numbers that can be used to classify textures. There are 14 textural features used to classify images according to Haralick. But only 5 features were used in this study. Namely ASM, contrast, IDM, entropy, and correlation. Figure 3 illustrates the steps in extracting texture features using the GLCM method.

Angular second moment (ASM) is used to measure uniformity or often called angular second moment. ASM will be high value when pixel values are similar to each other or in other words, the pixel distribution is in a constant condition. The ASM value is calculated from the

sum of squares of the coefficient matrix as shown in equation (1).

$$ASM = \sum_i \sum_j \{p(i, j)\}^2 \quad (1)$$

Contrast contains information that shows the size of the spread (moment of inertia) elements of the image matrix. If it is located far from the main diagonal, the contrast value is large. Visually value contrast is a measure of variation between the gray degrees of an image area. The value can be calculated as shown in equation (2).

$$Contrast = \sum_i \sum_j |i - j|^2 p(i, j) \quad (2)$$

Inverse Difference Moment (IDM) shows homogeneity of images of similar degree. Equation (3) is an equation for calculating the value of IDM.

$$IDM = \sum_i \sum_j \frac{1}{1 + (i - j)^2} p(i, j) \quad (3)$$

Entropy shows the size of the non-rule of an image. The value of an entropy is large if the image has an even gray degree and the entropy value will be of small value if the image structure is irregular (varied). Equation (4) is an equation to calculate the entropy value.

$$Entropy = - \sum_i \sum_j p(i, j) \log_2 p(i, j) \quad (4)$$

Correlation is a feature that shows a measure of linear dependence on the gray level of an image so that it can provide an indication of the existence of a linear structure (the joint probability). To calculate the value of the correlation shown by equations (5) to (9).

$$Correlation = \sum_i \sum_j \frac{(i, j)p(i, j) - \mu_i \mu_j}{\sigma_i \sigma_j} \quad (5)$$

Where

$$\mu_i = \sum_i i \sum_j p(i, j) \quad (6)$$

$$\mu_j = \sum_j j \sum_i p(i, j) \quad (7)$$

$$\sigma_i^2 = \sum_i \sum_j p(i, j)(i - \mu_i)^2 \quad (8)$$

$$\sigma_j^2 = \sum_i \sum_j p(i, j)(j - \mu_j)^2 \quad (9)$$

Notes :

$p(i, j)$ = is the value in row i and column j in the co-occurrence matrix.

μ_i and μ_j = is the mean value that shows the size of the dispersion of an image.

σ_i and σ_j = is the standard deviation which is a measure that describes the level spread of data from average values.

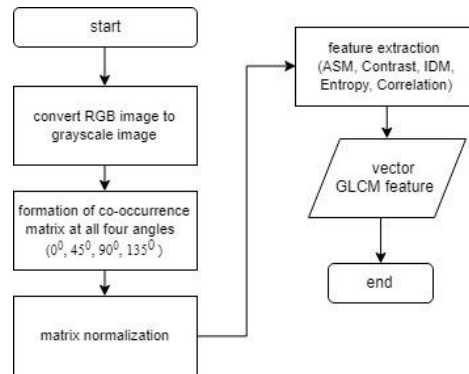


Figure 3 The steps of extracting features GLCM

2. 2.2 Hue, Saturation, Value (HSV)

The HSV color space is more intuitive to how people experience color than the RGB color space. As hue (H) varies from 0 to 1.0, the corresponding colors vary from red, through yellow, green, cyan, blue, and magenta, back to red. As saturation(S) varies from 0 to 1.0, the corresponding colors (hues) vary from unsaturated (shades of gray) to fully saturated (no white component). As value (V), or brightness, varies from 0 to 1.0, the corresponding colors become increasingly brighter. The hue component in HSV is in the range 0° to 360° angle all lying around a hexagon as shown in figure 4 HSV color space [8].

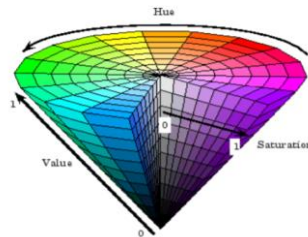


Figure 4 HSV color space

To get the value H, S, V based on R, G, B is shown in the equation (10) to (14) [9]. Conversion to HSV values is performed on each pixel of an RGB image. After that, the average is searched for H (mean H), S (mean S) and V (mean V). This average value will characterize the color feature. Figure 5 illustrates the steps to convert RGB to HSV and calculate mean H, mean S, mean V.

$$r = \frac{R}{R+G+B}, g = \frac{G}{R+G+B}, b = \frac{B}{R+G+B} \quad (10)$$

$$V = \max(r, g, b) \quad (11)$$

$$S = \begin{cases} 0 & , \text{ if } V = 0 \\ 1 - \frac{\min(r, g, b)}{V} & , \text{ if } V > 0 \end{cases} \quad (12)$$

$$\begin{aligned}
 & 0, \text{ if } S = 0 \\
 & \frac{60*(g-b)}{S*V}, \text{ if } V = r \\
 H = & 60*[2 + \frac{(b-r)}{S*V}], \text{ if } V = g \\
 & 60*[4 + \frac{(r-g)}{S*V}], \text{ if } V = b
 \end{aligned}
 \tag{13}$$

$$H = H + 360, \text{ if } H < 0
 \tag{14}$$

Notes :

- r, g, b = is normalization of the values of R (red), G (green), B (blue).
- R, G, B = is the values of R (red), G (green), B (blue).
- H, S, V = is the value of the conversion of R, G, B to H, S, V.

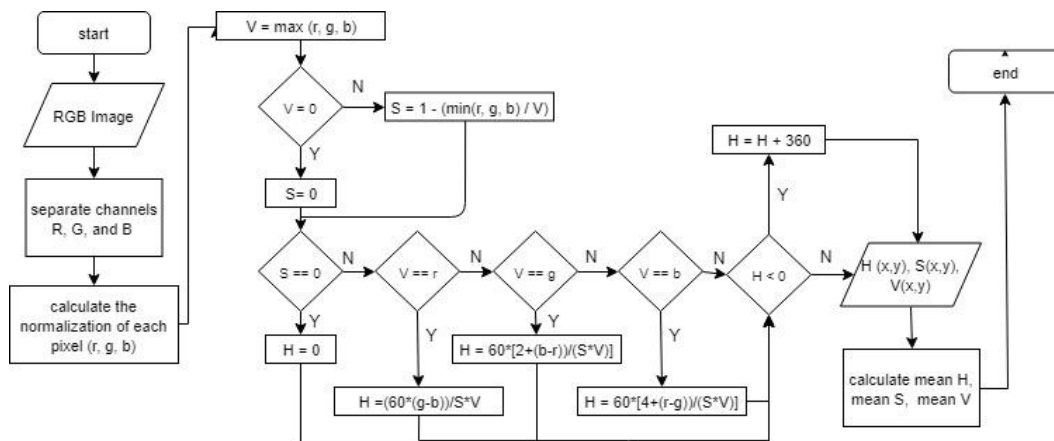


Figure 5 The steps of convert RGB to HSV and calculate mean H, mean S, mean V

2.3 4-fold Cross – Validation

From a total of 120 images will be divided into training data and test data using the 4-fold cross-validation method. In this method, all data is divided into 4 groups equally. Where 1 group will be the test data while the other 3 will be the training data [10]. From these 4 combinations of training and test data, verification results are sought for use in other testing scenarios. Figure 6 shows the distribution of training data and test data by group. So that each group has 30 data as test data and 90 data as training data.

fold 1		fold 2		fold 3		fold 4	
Keprok busuk	Keprok tidak	Keprok busuk	Keprok tidak	Keprok busuk	Keprok tidak	Keprok busuk	Keprok tidak
15	15	15	15	15	15	15	15
15	15	15	15	15	15	15	15
15	15	15	15	15	15	15	15
15	15	15	15	15	15	15	15

data uji (top row of each fold)
 data latih (bottom three rows of each fold)

Figure 6 The distribution of training data and test data by 4-fold cross – validation

2. 4 *k*-Nearest Neighbor (*k*-NN)

The *k*-Nearest Neighbor (*k*-NN) classification algorithm is one of the most popular approaches used by researchers and practitioners in the areas of Pattern Recognition and Machine Learning. Generally speaking, *k*-NN only needs one parameter to be adjusted, *k*, which represents how many closest neighbors are to be considered to classify an unseen object. Once this parameter is set, two main approaches are followed to classify an object, the vote of the majority of the *k* neighbors, and a weighted vote of all *k* neighbors considering the distance from where each of them is located concerning the object to classify [11] .

For image classification, a method is needed to calculate the distance between image training data and image test data. one of the most popular is the euclidean distance which will be discussed in 2.4.1. Figure 7 illustrates the steps of *k*-NN.

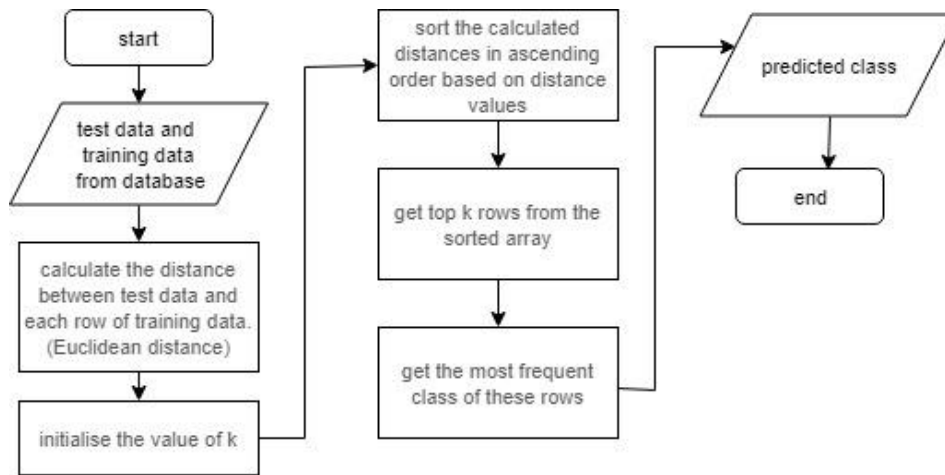


Figure 7 The steps of K-NN

2. 4.1 Euclidean distance

Euclidean distance is used to determine the level of similarity or inequality of two input vectors. The level of similarity is in the form of a score and based on the score two feature vectors will be said to be similar or not. Euclidean distance counts the root of the square of the difference of two vectors. Equation euclidean distance is shown in the equation (15).

$$d_{ij} = \sqrt{\sum_{k=1}^n (f_{ik} - f_{jk})^2} \quad (15)$$

Notes :

d_{ij} = is the Euclidean distance between vectors i and vector j.

f_{ik} = is the value of the k feature in the image i.

f_{jk} = is the value of the k feature in the image j.

2. 5 Confusion Matrix

A confusion matrix is a table that describes the performance of a classifier/classification model. It contains information about the actual and prediction classifications done by the classifier and this information is used to evaluate the performance of the classifier. To ascertain

the performance of a classifier using the confusion matrix and the data it contains, there are some standard metrics that we can calculate using the data(variables) in the confusion matrix. There are Accuracy, Precision, Sensitivity, and Specificity [12]. Table 1 shows the confusion matrix for the two-class classification problem.

Table 1 The confusion matrix for two-class classification problem

	Predicted Negative	Predicted Positive
Actual Negative	a	b
Actual Positive	c	d

a is the number of correct negative predictions, b is the number of incorrect positive predictions, c is the number of incorrect negative predictions, and d is the number of correct positive predictions. Accuracy is a percentage of correct predictions based on existing conditions. To calculate the Accuracy is shown in the equation (16).

$$Accuracy = \frac{a + d}{a + b + c + d} \quad (16)$$

Precision is calculating the proportion of a predictable class according to the correct conditions for all predicted data correctly. To calculate the Precision is shown in the equation (17).

$$Precision = \frac{a}{a + c} \quad (17)$$

Sensitivity (Recall / True positive rate) is used to measure the proportion of a class that has been predicted to match according to actual conditions. To calculate the Sensitivity is shown in the equation (18).

$$Sensitivity = \frac{a}{a + b} \quad (18)$$

Specificity (True negative rate) is the proportion of a class that is correctly predicted not from that class. To calculate the Specificity is shown in the equation (19).

$$Specificity = \frac{d}{c + d} \quad (19)$$

3. RESULTS AND DISCUSSION

In this study, using several test scenarios to see which scenarios can provide the highest accuracy value. First of all, determine the division of training data and test data using 4-fold cross-validation. Division of data folds can be seen in Figure 6. To test each data fold used a combination of GLCM and HSV feature extraction and classification using k-NN with a value of k 3.5, and 7. The performance of each data fold can be seen in Table 2.

Table 2 Average 4-fold cross-validation test performance results

Average 4-fold cross-validation test performance results				
	Accuracy	Precision	Sensitivity	Specificity
Fold 1	46%	46%	42%	49%
Fold 2	54%	60%	29%	80%
Fold 3	71%	75%	64%	78%
Fold 4	73%	81%	60%	87%

From the data shown in table 2, the best accuracy performance is shown in data fold 4. Then, for other test scenarios, it will use data fold 4. The next test is to compare the system performance for the extraction features of GLCM, HSV, and a combination of GLCM and HSV. The classification uses the k-NN method with k values 3, 5 and 7. Besides that, it is also tested for test data that does not have data labels (unlabeled).

Table 3 Performance test results for the classification of tangerines

		Accuracy	Precision	Sensitivity	Specificity
GLCM					
labeled test data	k = 3	57%	63%	33%	80%
	k = 5	67%	69%	60%	73%
	k = 7	73%	73%	73%	73%
unlabeled test data	k = 3	57%	60%	40%	73%
	k = 5	57%	58%	47%	67%
	k = 7	53%	56%	33%	73%
HSV					
labeled test data	k = 3	63%	83%	33%	93%
	k = 5	67%	86%	40%	93%
	k = 7	63%	83%	33%	93%
unlabeled test data	k = 3	87%	92%	80%	93%
	k = 5	93%	100%	87%	100%
	k = 7	93%	100%	87%	100%
GLCM & HSV					
labeled test data	k = 3	63%	75%	40%	87%
	k = 5	80%	91%	67%	93%
	k = 7	77%	79%	73%	80%
unlabeled test data	k = 3	63%	67%	53%	73%
	k = 5	70%	71%	67%	73%
	k = 7	67%	69%	60%	73%

Table 3 shows performance test results for the classification of tangerines. It can be seen that the highest performance value is obtained from the HSV feature extraction with unlabeled test data at values k = 5 and 7 that is 93%. Even though the value of the feature extraction performance by the combined method is lower than HSV, we still need to test the combined extraction method. Why is it necessary to combine the GLCM and HSV methods for features

extraction is because these two methods extract different features. GLCM for texture features while HSV for color features. the object of this study was tangerine which would be classified into a class of rotten tangerines and tangerines were not rotten. in visual observation, a clear difference is a change in color, shape, and texture.

This system has a high sensitivity performance percentage of 87% for testing HSV feature extraction with unlabeled test data at values $k = 5$ and 7 . In specificity performance, this performance is calculated to measure how well the training image in the class so that it does not affect other classes. Specificity is the opposite of sensitivity. The highest specificity performance was found in HSV feature extraction testing with unlabeled test data at $k = 5$ and 7 that is 100%. Precision Performance is used to calculate the proportion or ratio of true positive predictions compared to overall positive predicted results. The highest precision performance is found in the HSV feature extraction test with the unlabeled test data at $k = 5$ and 7 that is 100%. This shows that the class of unrotten tangerine does not affect the results of the classification of a rotten tangerine class.

Accuracy Performance is the correct percentage of data predicted in a class against all test data. The highest accuracy in this system is shown in the HSV feature extraction test with the unlabeled test data at $k = 5$ and 7 that is 93%. This shows that the amount of incorrect classification data is small.

4. CONCLUSIONS

The system can classify the quality of tangerines which are rotten and not rotten. The use of 4-fold cross-validation can help find the right combination of training data and test data to get good system performance values. Combined GLCM and HSV feature extraction methods provide the best accuracy performance value of 80% in tests with a value of $k = 5$ using labeled test data . While the best accuracy performance value of 93% is obtained from the HSV feature extraction tests with values of $k = 5$ and 7 on k-NN using unlabeled test data.

5. SUGGESTION

Rotten tangerine has the highest sensitivity performance of 100% in this test. there needs to be further study by combining this method so that the other types can also be classified properly. In this study, only one orange variety was used (tangerine). for future work, this combination method can be used to classify the quality of more fruit varieties.

REFERENCES

- [1] K. P. D. J. Hortikultura, *Statistik Pertanian 2018*. Jakarta: Pusat Data dan Sistem Informasi Pertanian, Kementrian Pertanian Republik Indonesia, 2018. [Online]. Available: <http://epublikasi.setjen.pertanian.go.id/download/file/438-statistik-pertanian-2018> [Accessed: 10-Aug-2019]
- [2] R. Widodo, A. W. Widodo, and A. Supriyanto, "Pemanfaatan Ciri Gray Level Co-Occurrence Matrix (GLCM) Citra Buah Jeruk Keprok (Citrus reticulata Blanco) untuk Klasifikasi Mutu," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 2, no. 11, pp. 5769–5776, 2018. [Online]. Available: <http://j-ptiik.ub.ac.id/index.php/j-ptiik/article/view/3420/1336> [Accessed: 16-Jul-2019]

- [3] C. Paramita, E. Hari Rachmawanto, C. Atika Sari, and M. S. Ignatius, "Klasifikasi Jeruk Nipis Terhadap Tingkat Kematangan Buah Berdasarkan Fitur Warna Menggunakan K-Nearest Neighbor," *J. Inform. J. Pengemb. IT*, vol. 4, no. 1, pp. 1–6, 2019.
- [4] O. R. Indriani, E. J. Kusuma, C. A. Sari, E. H. Rachmawanto, and D. R. Setiadi, "Tomatoes classification using K-NN based on GLCM and HSV color space," *Proc. - 2017 Int. Conf. Innov. Creat. Inf. Technol. Comput. Intell. IoT, ICITech 2017*, vol. 2018-Janua, no. November, pp. 1–6, 2018. [Online]. Available: https://www.researchgate.net/publication/323861268_Tomatoes_classification_using_K-NN_based_on_GLCM_and_HSV_color_space [Accessed: 16-Jul-2019]
- [5] 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. [Online]. Available: https://jurnal.ugm.ac.id/ijccs/article/view/17416/16693* [Accessed: 16-Jul-2019]
- [6] Karimah F.U., *Klasifikasi Citra Batik Kain Besurek Dengan Speed Up Robust Features (Surf) Dan Gray Level Co-Occurrence Matrix (GlcM)*. Yogyakarta, 2017.
- [7] R. A. Rasool, "Iris Feature Extraction and Recognition based on Gray Level Co-occurrence Matrix (GLCM) Technique," *Int. J. Comput. Appl. (0975 – 8887)*, vol. Volume 181, no. January, 2018. [Online]. Available: https://www.researchgate.net/publication/329718619_Iris_Feature_Extraction_and_Reco gnition_based_on_Gray_Level_Co-occurrence_Matrix_GLCM_Technique [Accessed: 17-Jul-2019]
- [8] S. Kolkur, D. Kalbande, P. Shimpi, C. Bapat, and J. Jatakia, "Human Skin Detection Using RGB, HSV and YCbCr Color Models," *ICCASP/ICMMD-2016. Adv. Intell. Syst. Res.*, vol. 137, pp. 324–332, 2017. [Online]. Available: <https://arxiv.org/ftp/arxiv/papers/1708/1708.02694.pdf> [Accessed: 17-Jul-2019]
- [9] A. Z. Maula, C. Rahmad, and U. D. Rosiani, "Pengembangan Aplikasi Pemilihan Buah Tomat Untuk Bibit Unggul Berdasarkan Warna Dan Ukuran Menggunakan HSV Dan Thresholding," *J. Teknol. Inf.*, vol. 7, no. 2, pp. 127–138, 2016. [Online]. Available: <http://ejurnal.stimata.ac.id/index.php/TI/article/view/223/253> [Accessed: 18-Jul-2019]
- [10] A. Wada *et al.*, "Differentiating Alzheimer's Disease from Dementia with Lewy Bodies Using a Deep Learning Technique Based on Structural Brain Connectivity," *Magn. Reson. Med. Sci.*, vol. 18, no. 3, pp. 219–224, 2019. [Online]. Available: [https://www.jstage.jst.go.jp/article/mrms/advpub/0/advpub_mp.2018-0091/_ pdf/-char/en](https://www.jstage.jst.go.jp/article/mrms/advpub/0/advpub_mp.2018-0091/_pdf/-char/en) [Accessed: 18-Aug-2019]
- [11] J. Aguilera, L. C. González, M. Montes-y-Gómez, and P. Rosso, "A new weighted k-nearest neighbor algorithm based on newton's gravitational force," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 11401 LNCS, no. November, pp. 305–313, 2019. [Online]. Available: https://www.researchgate.net/publication/329129843_A_New_Weighted_k-Nearest_Neighbor_Algorithm_based_on_Newton's_Gravitational_Force [Accessed: 18-Jul-2019]

- [12] Akhyar, R. Hidayat, and B. Hantono, "Identifikasi Fitur Wajah Dengan Menggunakan Facial Points Dan Euclidean Distance," *Citee*, pp. 154–160, 2017.