Journal of Computer Networks, Architecture and High Performance Computing
Volume 5, Number 1, January 2023
https://doi.org/10.47709/cnahpc.v5i1.1985

Submitted : January 6, 2023
Accepted : January 14, 2023
Published : January 15, 2023

Classification of Banana Ripeness Based on Color and Texture Characteristics

Ahmad Hafidzul Kahfi[1)], Muhamad Hasan[2)], Riyan Latifahul Hasanah[3)]
[1)2)3)]Nusa Mandiri University, Indonesia
[1)]ahmad.azx@nusamandiri.ac.id, [2)]muhamad.mhx@nusamandiri.ac.id, [3)]riyan.rlt@nusamandiri.ac.id

ABSTRACT

Banana is one of the most consumed fruits globally and is a rich source of vitamins, minerals and carbohydrates. With the many benefits that bananas have, many farmers cultivate this fruit. The problem that occurs when the harvest is produced on a large scale is the process of selecting bananas that are still unripe or ripe. Usually farmers carry out the selection process manually by visually identifying ripeness based on the color of the fruit skin. However, direct observation has several drawbacks such as subjectivity, takes a long time and is inaccurate. For this reason, we need a system that can help determine the maturity level of bananas automatically through a series of banana image processing processes. One way that can be used to determine the maturity level of bananas is by looking at the color and texture of the bananas. This study aims to classify the maturity level of bananas based on the color and texture characteristics of the banana image using the Gray Level Co-occurrence Matrix and K-Nearest Neighbor methods for the classification process. Based on the results of the research analysis that has been carried out, using the parameter k which has a value of 3 obtains very high accuracy.

Keywords: Classification , Fruit Maturity, Gray Level Co-Occurrence Matrix, K-Means Clustering, K-Nearest Neighbor

INTRODUCTION

Bananas are one of the most consumed fruits globally (Mazen & Nashat, 2019) and the fourth most important food crop in the world along with rice, wheat and corn (Surya Prabha & Satheesh Kumar, 2015), in terms of gross domestic product, with world production of around 70 million tons in 2003 (UNCTAD 2003) (Jaiswal et al., 2014).

This contributes around 16% of world fruit production according to FAO (Mazen & Nashat, 2019). Banana is also a staple commodity that is very important for many developing countries, so it has relevance for its role in food security (Jaiswal et al., 2014). Besides being healthy and delicious bananas also contain several important nutrients, and have benefits for digestion, heart health, and weight loss as well as one of the high calories of tropical fruits (Toma et al., 2018). Because of the many benefits that cause bananas are cultivated.

Bananas are usually harvested in the mature green stage and they remain hard and green without significant changes in skin color, texture, or composition (depending on temperature, humidity, and age of the banana at harvest) before the onset of ripening (Dadzie and Orchard 1997) (Mendoza & Aguilera, 2006). The maturity stage when harvesting fresh bananas is an important factor affecting fruit quality during marketing (Toma et al., 2018).

Fruit ripening has become an important concern and is categorized as climacteric (ripe after harvest) and non-climacteric (not ripe from trees) (Marimuthu & Mohamed Mansoor Roomi, 2017). Because bananas can be planted in many countries and produced throughout the year (Dittakan et al., 2017), bananas have become the fourth largest climacteric commodity (Marimuthu & Mohamed Mansoor Roomi, 2017) which can be used for domestic consumption and export (Dittakan et al., 2017).

Bananas are usually evaluated based on their skin color (Kipli et al., 2018). Color is a very important fruit quality characteristic. This represents the level of maturity, sugar content, acidity, and taste (Wan et al., 2018). Skin color is considered as the first quality parameter evaluated by consumers (Mendoza & Aguilera, 2006) and one of the benchmarks of customers to buy bananas from the market is their skin color (Sabilla et al., 2019). According to (Saputro et al., 2018), changes in skin color can indicate the maturity of bananas, which are initially green to yellow due to the degradation process of chlorophyll pigments.

Farmers usually determine the level of maturity of bananas by looking at the characteristics of bananas through direct observation. However, direct observation has a high error rate (Kaur et al., 2018), is very subjective, time-

* Corresponding author

This is an Creative Commons License This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International (CC BY-SA 4.0).

Journal of Computer Networks, Architecture and High Performance Computing
Volume 5, Number 1, January 2023
https://doi.org/10.47709/cnahpc.v5i1.1985

Submitted : 6 January 2023
Accepted : 14 January 2023
Published : 15 January 2023

consuming and tiring (Mustaffa & Khairul, 2018). For that, we need a system that can determine the level of maturity of bananas automatically through a series of image processing processes of bananas. One way that can be used to determine the maturity level of bananas is by looking at the color and texture of the banana. One method of texture analysis that can be used is the Gray Level Co-occurrence Matrix method and RGB space transformation for color feature extraction. To analyze the characteristics of color and texture must go through image processing by utilizing digital image processing techniques.

From the extraction process, the color and texture features produce values that are used as input to the classification process. One classification method that can be used is K-Nearest Neighbor, which is a method that is robust against noise training data and is effective when the data is large. This study proposes the classification of the maturity level of bananas based on color and texture characteristics. It is expected that the results of this study will be able to classify the level of ripeness of bananas.

LITERATURE REVIEW

Image processing

Digital image processing is generally defined as two-dimensional image processing with a computer (Wibisono & Mujiyono, 2018) to obtain more perfect results so that the information contained in the image can be seen more clearly (Suta et al., 2019). The image processing process begins with input in the form of images and produces output in the form of processed images (Ratnawati & Sulistyaningrum, 2019).

Image Segmentation

Image segmentation is an image processing process that aims to separate the object region from the background region so that objects are easily analyzed in order to identify objects that involve a lot of visual perception (Ratnawati & Sulistyaningrum, 2019) where each region has similar attributes (Andika & Hafiz, 2018).

K-Means

The K-Means algorithm is one of the partitional algorithms, because K-Means is based on determining the initial number of groups by defining the initial centroid value (Andika & Hafiz, 2018). Simply put, K-Means is an algorithm for classifying attributes into k groups, where k is a positive integer. Grouping is done by minimizing the distance between data and based on centralized grouping. The distance used in K-means Clustering is the euclidean distance. Euclidean distance is the distance between two points in n-dimensional space by calculating the root of the difference between the two points or paired objects.

Feature Extraction

The main purpose of feature extraction is to characterize the object you want to identify from an image by using measures that have very similar values (Andika & Hafiz, 2018) to obtain the most relevant information from the original data and represent that information in a lower dimensional space (Ratnawati & Sulistyaningrum, 2019). Shape features, geometry features, size features, texture features and color features are features that are often extracted. Where to get these features there are differences in each feature extracted. One of them is the size feature which has several parameters such as area or area which is the number of pixels that make up an object. While the perimeter or circumference is the number of pixels that are at the object boundary. This feature is then used by the classifier to identify the input unit with the target output unit. The classifier makes it easier to classify between different classes by looking at the existing features (Setiawan & Putra, 2018). The features extracted in this study are color features with the average parameters of the Red, Green and Blue channels as well as texture features using GLCM with parameters of contrast, correlation, energy and homogeneity.

Gray Level Co-occurrence Matrix Method

The Gray Level Co-occurrence Matrix (GLCM) method is a method that analyzes a pixel in an image and determines the gray level that often occurs (Andika & Hafiz, 2018) and produces a co-occurrence matrix whose elements are the number of occurrences of pixels that have a level value. certain gray area, where the pixel pairs are at a certain distance (d) and angle (Θ) (Pamungkas, 2019). Distance is represented as pixels whereas orientation is represented in degrees. The orientation is formed from four angular directions with intervals of 45, namely 0, 45, 90 and 135. To perform image analysis based on the statistical distribution of the pixel intensity, it can be done by

* Corresponding author

This is an Creative Commons License This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International (CC BY-SA 4.0).

Journal of Computer Networks, Architecture and High Performance Computing
Volume 5, Number 1, January 2023
https://doi.org/10.47709/cnahpc.v5i1.1985

Submitted : 6 January 2023
Accepted : 14 January 2023
Published : 15 January 2023

extracting the texture features. GLCM is a method for performing statistical-based feature extraction, feature acquisition is obtained from matrix pixel values, which have a certain value and form a pattern angle.

K-Nearest Neighbor

K-Nearest Neighbor is an approach to find data by calculating the closeness between new data and old data (Pamungkas, 2019). New data can be classified based on the majority group of k nearest neighbors (Febrinanto et al., 2018), where k is a positive integer and usually a small number. Classification is carried out on objects based on learning data that are closest to the object. K-Nearest Neighbor is an instance-based classification method, selecting a training object that has the closest neighbor properties. This neighborly property is obtained from calculating the value of similarity or dissimilarity. KNN uses the method of calculating dissimilarity values (Euclidian, Manhattan, Square Euclidian, etc.). KNN will choose the k nearest neighbors to determine the classification results by looking at the number of occurrences of the class in the selected k neighbors. The class that appears the most will be the class resulting from the classification.

METHOD

The steps undertaken in this study are shown in Figure 1.

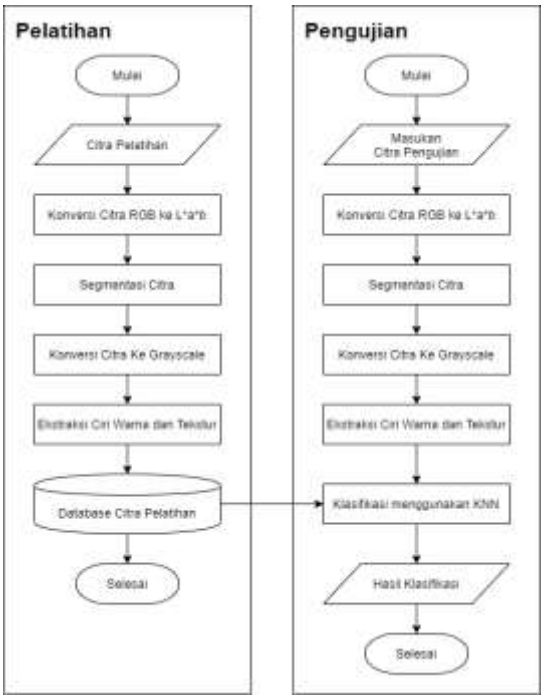


Fig 1. Research Stages

Figure 1 is the flow of research conducted with the following discussion:

Training Image

In inputting training images, training image datasets that have been acquired previously are used. The dataset was acquired from bananas which have 2 levels. The training image dataset acquired 144 images, measuring 720 x 720 pixels and acquired during the day which is divided into 2 classes with each class totaling 72 images. The image was acquired using a Vivo Y93 smartphone camera with a resolution of 13 MP with a distance of 20 cm from the banana to the camera.

* Corresponding author

This is an Creative Commons License This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International (CC BY-SA 4.0).

Journal of Computer Networks, Architecture and High Performance Computing
Volume 5, Number 1, January 2023
https://doi.org/10.47709/cnahpc.v5i1.1985

Submitted : 6 January 2023
Accepted : 14 January 2023
Published : 15 January 2023

Convert RGB Image to L*a*b

After the training image is inputted, the next step is to change the image format from RGB to L*a*b. This is done in order to facilitate the segmentation process in the next stage.

Image Segmentation

Image that has been converted into L*a*b format will be segmented to separate the object to be extracted from other objects that are not needed.

Convert Image to Grayscale

After obtaining the image from the segmentation results, the next step is to change the format of the image segmentation results in grayscale images. This is done because the texture feature extraction process using the GLCM method must be done on grayscale images.

Color and Texture Feature Extraction

From the results of the conversion of the grayscale image in the previous stage, it can be done the process of extracting the color and texture characteristics of the image of a banana. The color features extracted are the image of R, G and B. channels, while the texture features extracted are contrast, correlation, energy, and homogeneity. To get 4 texture features to use the following formula:

Energy is a measure of the homogeneity of the image. Energy shows a high value when the pixels of the image are homogeneous.

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

Contrast is a measure of the existence of variations in the level of gray one pixel with adjacent pixels throughout the image.

$$CON = \sum_k k^2 \left[\sum_i \sum_j p(i,j) \right] \qquad (2)$$

Where:

$|i-j| = k$

Correlation shows a measure of the linearness of the gray level of one pixel relative to other pixels at a certain position.

$$COR = \frac{\sum_i \sum_j (ij).p(i,j)\mu_x\mu_y}{\sigma_x\sigma_y} \qquad (3)$$

Homogeneity measures the proximity of the element distribution in GLCM to diagonal GLCM.

$$HOM = \sum_i \sum_j \frac{p(i,j)}{1+|i-j|} \qquad (4)$$

The texture feature extraction process that has been done produces the value of each feature made into the database to form rows and columns.

Training Image Database

From the previous stage, the values of each feature are generated, namely R, G, B, contrast, correlation, energy, and homogeneity. The value is stored in the form of rows and columns so that a training image database is formed to be used again at the testing stage.

* Corresponding author

This is an Creative Commons License This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International (CC BY-SA 4.0).

Journal of Computer Networks, Architecture and High Performance Computing
Volume 5, Number 1, January 2023
https://doi.org/10.47709/cnahpc.v5i1.1985

Submitted : 6 January 2023
Accepted : 14 January 2023
Published : 15 January 2023

DISCUSSION

The first stage in the study was the creation of a database for training images using 144 images. This database will be used in the testing phase by calculating the distance between the value of the training image and the test image. The database creation results can be seen in Table 1.

Table 1. Database of Training Imagery

No	Red	Green	Blue	Contrast	Correlation	Energy	Homogeneity
1	165,9604288	121,3555744	72,05279312	0,060263128	0,98357952	0,713490808	0,991522434
2	182,5579015	137,4861317	79,45289608	0,079056283	0,984852013	0,669001003	0,991183531
3	183,2222839	136,8930204	74,23939018	0,090989696	0,985094426	0,605611117	0,987917295
…	…	…	…	…	…	…	…
142	110,7048884	111,3150125	63,54412563	0,088714719	0,96919725	0,650799505	0,985507222
143	110,2392969	106,3076741	63,12229468	0,074156899	0,964648874	0,705975655	0,986597803
144	109,7229854	112,405547	56,3093536	0,059887879	0,977013889	0,662779063	0,988349542

In this study, testing was carried out by processing 36 images of bananas with JPG extension. Of the 36 images will be processed based on color and texture with 18 images in each class that is ripe and raw. The results of image capture can be seen in Figure 2.



Fig 2. Testing Image

The test image is processed first by converting it from an RGB image to an L * a * b image. The results of the image conversion process can be seen in Figure 3.



Fig 3. Converting an RGB Image to L * a * b

* Corresponding author

This is an Creative Commons License This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International (CC BY-SA 4.0).

Journal of Computer Networks, Architecture and High Performance Computing
Volume 5, Number 1, January 2023
https://doi.org/10.47709/cnahpc.v5i1.1985

Submitted : 6 January 2023
Accepted : 14 January 2023
Published : 15 January 2023

The next step is to segment the image from the L * a * b image to separate the object to be processed with other objects. The results of segmentation can be seen in Figure 4.



Fig 4. Image Segmentation using K-Means Clustering from L * a * b images

After getting the segmented image, the next step changes the segmented image to a grayscale image so that the texture extraction process can be done using the GLCM method. The results of image conversion can be seen in Figure 5.

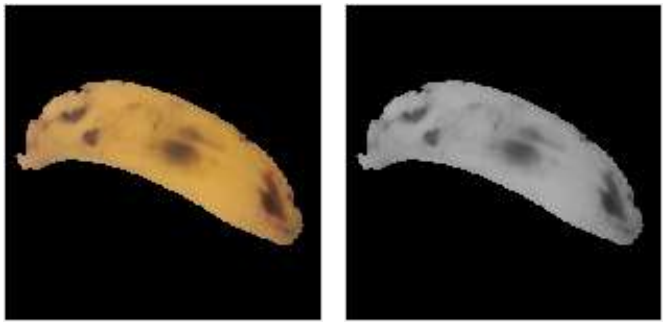


Fig 5. The Segmented Image is converted to Grayscale Image

After getting a grayscale image, the next step can be done, namely the extraction of color and texture features. Where the value of these features will be used for the identification process. The results of the feature value extraction process can be seen in Figure 6.

Hasil Ekstraksi Fitur

	Fitur	Nilai
1	Red	183.2223
2	Green	136.893
3	Blue	74.2394
4	Contrast	0.09099
5	Correlation	0.98509
6	Energy	0.60561
7	Homogeneity	0.98792

Fig 6. Color and Texture Extraction Results

* Corresponding author

This is an Creative Commons License This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International (CC BY-SA 4.0).

Journal of Computer Networks, Architecture and High Performance Computing
Volume 5, Number 1, January 2023
https://doi.org/10.47709/cnahpc.v5i1.1985

Submitted : 6 January 2023
Accepted : 14 January 2023
Published : 15 January 2023

To see the program display in MATLAB can be seen in Figure 7.

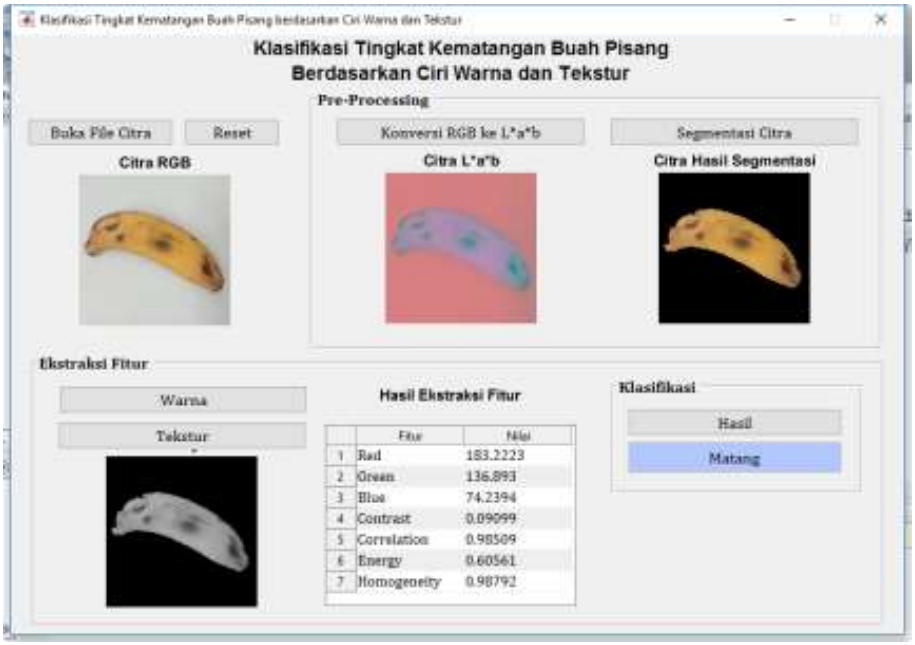


Fig 7. GUI display in MATLAB

RESULT

In this study, testing with a parameter k value of 3. The results obtained in this study can be seen in Table 2.

Table 2. Testing Results

No	Classification	Test result		Accuracy
		True	False	
1	Raw	16	2	89%
2	Ripe	15	3	83%
Average				86%

CONCLUSION

Based on the research that has been done, it can be concluded that the classification of the ripeness level of bananas can be done using digital image processing based on color and texture. Through the KNN (K-Nearest Neighbor) algorithm, results are obtained with very high accuracy with a value of k parameter 3.

In addition, digital image extraction based on color and texture by extracting Energy, Contrast, Correlation, and Homogeneity values can increase accuracy in carrying out the classification process.

In the future, this research can be developed using other types of classifications, for example, Naive Bayes classification or Support Vector Machine and others. As well as adding other feature extractions and adding several classes to the level of fruit maturity.

REFERENCES

Andika, T. H., & Hafiz, A. (2018). Analisis Perbandingan Segmentasi Citra Menggunakan Metode K-Means dan Fuzzy C-Means. *Seminar Nasional Teknologi Dan Bisnis 2018*, 237–246.

Dittakan, K., Theera-Ampornpunt, N., Witthayarat, W., Hinnoy, S., Klaiwan, S., & Pratheep, T. (2017). Banana Cultivar Classification using Scale Invariant Shape Analysis. *Proceeding of 2017 2nd International*

* Corresponding author

This is an Creative Commons License This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International (CC BY-SA 4.0).

Journal of Computer Networks, Architecture and High Performance Computing
Volume 5, Number 1, January 2023
https://doi.org/10.47709/cnahpc.v5i1.1985

Submitted : 6 January 2023
Accepted : 14 January 2023
Published : 15 January 2023

Conference on Information Technology, INCIT 2017, 1–6. https://doi.org/10.1109/INCIT.2017.8257854

Febrinanto, F. G., Dewi, C., & Wiratno, A. T. (2018). Implementasi Algoritme K-Means Sebagai Metode Segmentasi Citra Dalam Identifikasi Penyakit Daun Jeruk. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer (J-PTIIK) Universitas Brawijaya*, *2*(11), 5375–5383.

Jaiswal, P., Jha, S. N., Kaur, P. P., Bhardwaj, R., Singh, A. K., & Wadhawan, V. (2014). Prediction of textural attributes using color values of banana (Musa sapientum) during ripening. *Journal of Food Science and Technology*, *51*(6), 1179–1184. https://doi.org/10.1007/s13197-012-0614-2

Kaur, H., Sawhney, B. K., & Jawandha, S. K. (2018). Evaluation of plum fruit maturity by image processing techniques. *Journal of Food Science and Technology*, *55*(8), 3008–3015. https://doi.org/10.1007/s13197-018-3220-0

Kipli, K., Zen, H., Sawawi, M., Noor, M. S. M., Julai, N., Junaidi, N., Razali, M. I. S. M., Chin, K. L., & Masra, S. M. W. (2018). Image Processing Mobile Application for Banana Ripeness Evaluation. *2018 International Conference on Computational Approach in Smart Systems Design and Applications, ICASSDA 2018*, 1–5. https://doi.org/10.1109/ICASSDA.2018.8477600

Marimuthu, S., & Mohamed Mansoor Roomi, S. (2017). Particle Swarm Optimized Fuzzy Model for the Classification of Banana Ripeness. *IEEE Sensors Journal*, *17*(15), 4903–4915. https://doi.org/10.1109/JSEN.2017.2715222

Mazen, F. M. A., & Nashat, A. A. (2019). Ripeness Classification of Bananas Using an Artificial Neural Network. *Arabian Journal for Science and Engineering*, *44*(8), 6901–6910. https://doi.org/10.1007/s13369-018-03695-5

Mendoza, F., & Aguilera, J. M. (2006). Application of Image Analysis for Classification of Ripening Bananas. *Journal of Food Science*, *69*(9), E471–E477. https://doi.org/10.1111/j.1365-2621.2004.tb09932.x

Mustaffa, I. B., & Khairul, S. F. B. M. (2018). Identification of fruit size and maturity through fruit images using OpenCV-Python and Rasberry Pi. *Proceeding of 2017 International Conference on Robotics, Automation and Sciences, ICORAS 2017*, *2018-March*, 1–3. https://doi.org/10.1109/ICORAS.2017.8308068

Pamungkas, D. P. (2019). Ekstraksi Citra menggunakan Metode GLCM dan KNN untuk Indentifikasi Jenis Anggrek (Orchidaceae). *Innovation in Research of Informatics (INNOVATICS)*, *1*(2), 51–56.

Ratnawati, L., & Sulistyaningrum, D. R. (2019). *Penerapan Random Forest untuk Mengukur Tingkat Keparahan Penyakit pada Daun Apel*. *8*(2), A71–A77.

Sabilla, I. A., Wahyuni, C. S., Fatichah, C., & Herumurti, D. (2019). Determining banana types and ripeness from image using machine learning methods. *Proceeding - 2019 International Conference of Artificial Intelligence and Information Technology, ICAIIT 2019*, 407–412. https://doi.org/10.1109/ICAIIT.2019.8834490

Saputro, A. H., Juansyah, S. D., & Handayani, W. (2018). Banana (Musa sp.) maturity prediction system based on chlorophyll content using visible-NIR imaging. *2018 International Conference on Signals and Systems, ICSigSys 2018 - Proceedings*, 64–68. https://doi.org/10.1109/ICSIGSYS.2018.8373569

Setiawan, K. N., & Putra, I. M. S. (2018). Klasifikasi Citra Mammogram Menggunakan Metode K-Means, GLCM, dan Support Vector Machine (SVM). *Jurnal Ilmiah Merpati (Menara Penelitian Akademika Teknologi Informasi)*, *6*(1), 13–24. https://doi.org/10.24843/jim.2018.v06.i01.p02

Surya Prabha, D., & Satheesh Kumar, J. (2015). Assessment of banana fruit maturity by image processing technique. *Journal of Food Science and Technology*, *52*(3), 1316–1327. https://doi.org/10.1007/s13197-013-1188-3

Suta, I. B. L. M., Hartati, R. S., & Divayana, Y. (2019). Diagnosa Tumor Otak Berdasarkan Citra MRI (Magnetic Resonance Imaging). *Majalah Ilmiah Teknologi Elektro*, *18*(2), 149–154. https://doi.org/10.24843/mite.2019.v18i02.p01

Toma, F., Ahmmed, R., Hasan, M., Haque, M., Monju, M., & Surovi, M. (2018). Non-destructive maturity index of "Amritsagor" banana using RGB and HSV values. *Journal of the Bangladesh Agricultural University*, *16*(2), 293–302. https://doi.org/10.3329/jbau.v16i2.37985

Wan, P., Toudeshki, A., Tan, H., & Ehsani, R. (2018). A methodology for fresh tomato maturity detection using computer vision. *Computers and Electronics in Agriculture*, *146*(February 2017), 43–50. https://doi.org/10.1016/j.compag.2018.01.011

Wibisono, I. S., & Mujiyono, S. (2018). Segmentasi Fuzzy C-Means Untuk Membantu Identifikasi Kualitas Beras Berdasarkan Nilai Threshold , Warna Dan Ukuran. *Multimatrix*, *I*(1), 22–25.

* Corresponding author

This is an Creative Commons License This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International (CC BY-SA 4.0).