JURNAL UPI

Journal of Software Engineering, Information ar Communication Technology (SEICT)

Journal homepage: https://ejournal.upi.edu/index.php/SEICT

SEICT

Determination of Mango Fruit Maturity on the Tree Based on Digital Image Processing and Artificial Neural Networks

Aditia Sanjaya, Ichwanul Muslim Karo Karo*

Faculty of Mathematic and Science, Universitas Negeri Medan, Indonesia
Correspondence: E-mail: aditiasanjaya@mhs.unimed.ac.id

ABSTRACT

Until now, humans have determined the ripeness of mangoes on the tree by hand. Losses are caused by the insecurity of the human state and a misunderstanding of the maturity level of mangoes. In the future, a system that can detect the ripeness of mangoes on the tree will be required. This research provides a preliminary examination of the technology's implementation. The study created a computerized image processing method for determining the ripeness of mangoes on the tree. The neural network backpropagation algorithm was employed in this investigation. The feature extraction model employed in the image is a hybrid of the RBG and HSV models. The best accuracy level is 72%, with an 80:20 ratio of test data to training data.

ARTICLE INFO

Article History:
Submitted/Received 15 May 2023
First Revised 20 May 2023
Accepted 25 May 2023
First Available online 29 May 2023
Publication Date 01 Jun 2023

Keyword:
Backpropagation,
HSV,
Mango,
RGB.

1. INTRODUCTION

Mango is a seasonal fruit that is widely known by the Indonesian community. It originated from India and had many varieties that are popular among Indonesians. Mango trees thrive in lowland areas with a hot climate. However, there are also varieties that can grow in regions with elevations of up to 600 meters above sea level.

Probolinggo, Situbondo, and Pasuruan in East Java are among the largest mango-producing regencies in Java Island. According to the Ministry of Agriculture (2015), mango production in Indonesia increased from 2012 to 2014. In 2012, the production reached 76,547 tons, followed by 98,958 tons in 2013, and 102,820 tons in 2014 (Sanjaya and Rosadi, 2018).

In the case of mango ripeness, sometimes a mango can still taste sour even when it is fully ripe, and vice versa (Hossain et al., 2021). Therefore, there is a need for technology to determine the level of ripeness of mangoes (Kurniawan and Junaidi, 2022). According to the Webster dictionary, an image is a representation, resemblance, or imitation of an object. For example, a photograph represents a person captured by a camera, an X-ray image represents a person's body part, and so on (Shi et al., 2020). There is a lot of information that can be extracted from an image because it provides visual information (Adenugraha et al., 2022). In general, digital image processing is a process that aims to manipulate and analyze images with the help of computers.

Previous research conducted by (Javanmardi et al., 2021) focused on image processing to determine the ripeness of grapes on the vine by combining the Radial Symmetry Transform algorithm and k-Nearest Neighbors, which achieved an accuracy of 90.2%. Therefore, it is highly possible to use digital image processing methods to analyze the ripeness level of fruits other than grapes (Javanmardi et al., 2021). Hence, this research will attempt to apply digital image processing to determine the ripeness level of mangoes that are still on the tree.

2. RESEARCH METHODS

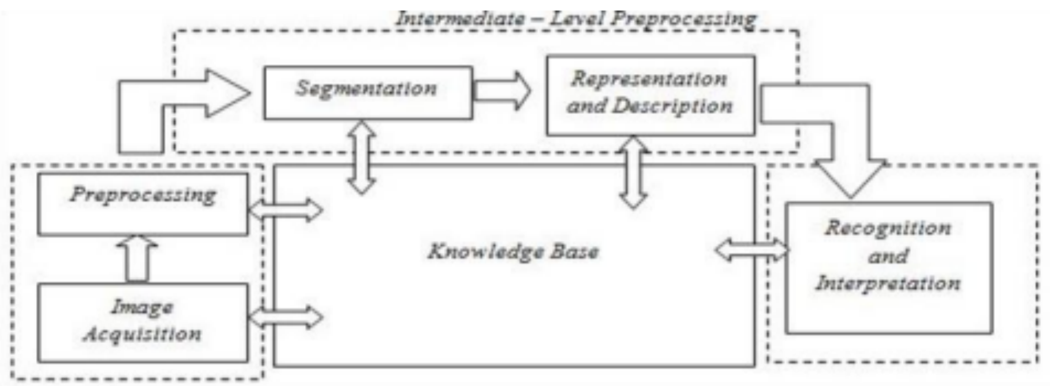

Figure 1. Key steps of digital image processing

2.1. Digital Image Processing

The first step in digital image processing is the image acquisition process, which involves capturing or obtaining the image. This can be done using sensors such as cameras, scanners,

DOI: https://doi.org/10.17509/seict.v4il.52916
p- ISSN 2774-1656 e- ISSN 2774-1699

and other devices. The next step is the preprocessing stage, where the image undergoes adjustments, such as resizing or enhancing its quality before being used for specific purposes (Zeb et al., 2007). The subsequent step is segmentation, which involves dividing the image into its constituent parts. This process aims to separate the desired objects from the rest of the image.

Further observation and analysis (representation and description) are then performed to verify that the regions within the boundaries represent the intended objects (Zhang and Lu, 2004). This step is necessary because the output of the segmentation process provides the boundaries between the object of interest and other objects in the image (Sutisna et al., 2020). The final step is recognition and interpretation. Recognition involves labeling the objects based on the information provided by their descriptors. Interpretation aims to extract meaningful information and understanding from the labeled objects. These steps collectively contribute to the analysis and interpretation of digital images as shown in **Figure 2a-2c** (Syarifah et al., 2022).

Figure 2a. An example image of a fruiting mango tree with raw

Figure 2b. An example image of a fruiting mango tree with medium

Figure 2c. An example image of a fruiting mango tree with ripe

DOI: https://doi.org/10.17509/seict.v4il.52916
p- ISSN 2774-1656 e- ISSN 2774-1699

2.2 Color Model and Color Features

In computer applications, colors are formed on the display interface using a specific color model. One widely used example of a color model is RGB (Alcaruban et al., 2018). The RGB color model belongs to the additive color model, where colors are formed by combining the light energy of three primary colors: red, green, and blue (Latuconsina, 2021). Different levels of intensity in each primary color component when combined will result in different colors (Nurraharjo, 2012).

The current computer screens typically have 256 levels of intensity (0-255) for each RGB component, allowing for a wide range of color gradations. In addition to the RGB color model, the HSV color model is also commonly used in digital image processing (Himmah et al., 2020). The HSV color model represents color using three components: hue (H), saturation (S), and value (V) (Salsabila et al., 2021). The HSV color model is a nonlinear model that approximates human visual color perception.

Features are distinctive characteristics that differentiate one object from another. In an image, there are features extracted from the patterns of colors, which can be used to differentiate and compare one image with another (Laia et al., 2023). Several color features can be derived from statistical characteristics of a group of RGB color component values or HSV color component values in an image (Alcahruban et al., 2018). Examples of such features include the mean values of the red, green, blue, hue, saturation, value, and grayscale components. In this study, the features used in the segmentation process are the mean values of R, G, B, and grayscale (Sari and Purnama, 2018). During the classification process (maturation determination), the features used are the mean values of R and V. These feature values are extracted from cropped images containing mango pixels. An example of a cropped image with its corresponding extracted feature values as shown in **Figure 3a-3c**.

"

Figure 3a. Examples of cropped images with their respective levels of raw

DOI: https://doi.org/10.17509/seict.v4il.52916
p- ISSN 2774-1656 e- ISSN 2774-1699

Figure 3b. Examples of cropped images with their respective levels of medium

Figure 3c. Examples of cropped images with their respective levels of ripe

2.3 Backpropagation Classification

Backpropagation is one of the algorithms used in artificial neural networks, specifically in supervised learning models (Chen and Zhong, 2009). The concept of backpropagation involves a multi-layer network that aims to minimize the error in the network's output (Paola and Schowengerdt, 1995). In the classification process using backpropagation, there are several steps involved. It starts with placing each object's features into the input layer neurons, corresponding to indices 1 and 2 (mean R and mean V). Then, an activation process is performed using a binary sigmoid activation function for each neuron in the hidden layer, followed by calculating the activation values for each neuron in the output layer. The calculated results are then used for thresholding. If the value is greater than or equal to 0.5, it is considered as 1. If the value is less than 0.5, it is considered as 0. Since there are only 2 neurons in the output layer, the possible output combinations [out1, out2] are [0,0] = ripe, [1,0] = medium, [0,1] = unripe, and any other combination is considered out of the system's

DOI: https://doi.org/10.17509/seict.v4il.52916
p- ISSN 2774-1656 e- ISSN 2774-1699

range. The flowchart of the maturity determination system using backpropagation as shown in **Figure 4**.

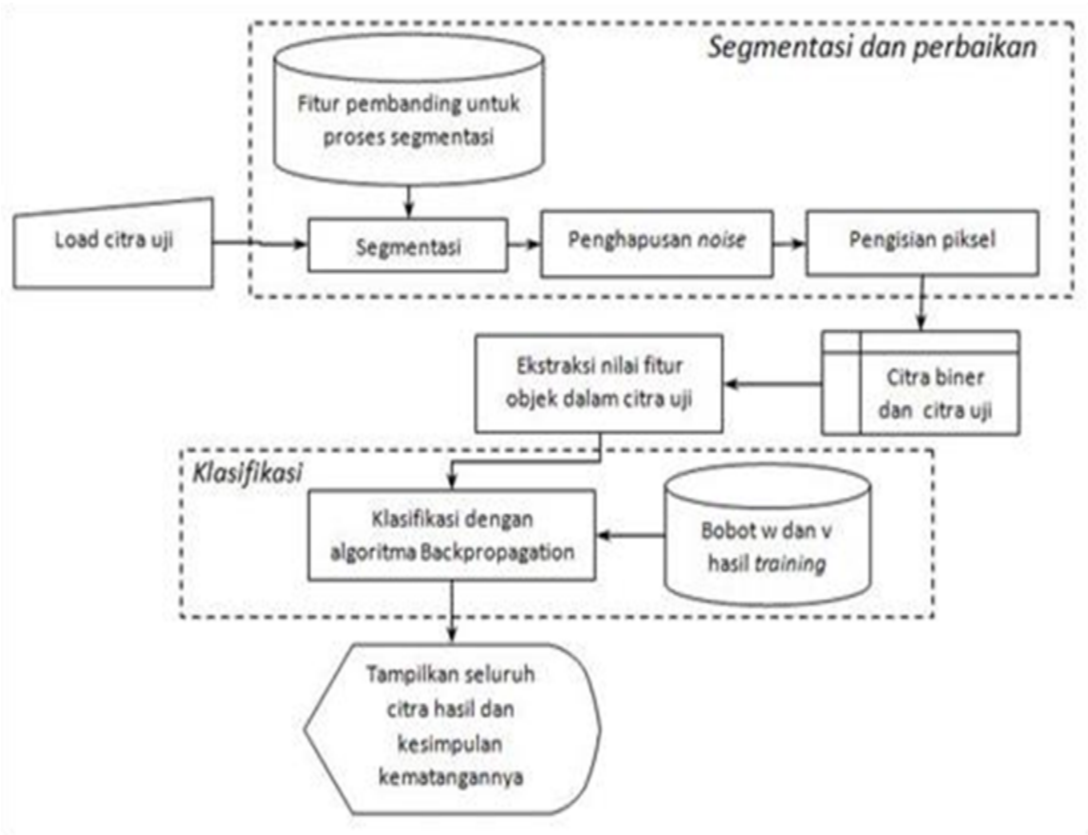

Figure 4. Flowchart of the process in the maturity determination system using Backpropagation

3. RESULTS AND DISCUSSION

To manage and develop a computer system for classifying the ripeness level of mangoes, a dataset is required to identify the existing issues. The dataset should consist of captured images of mangoes at different ripeness levels, ranging from unripe to overripe. In this research, a dataset of 52 mangoes with varying ripeness levels was used, including 11 unripe mangoes, 12 medium-ripe mangoes, 11 ripe mangoes, and 10 overripe mangoes. This data was then processed to be learned by the artificial neural network system.

Identifying the ripeness level of mangoes manually is actually easy to implement, as it can be visually observed from the color and through touch by assessing the mango's texture. However, implementing manual identification for a large quantity of mangoes is considered less effective as shown in **Figure 5a-5d**.

DOI: https://doi.org/10.17509/seict.v4il.52916
p- ISSN 2774-1656 e- ISSN 2774-1699

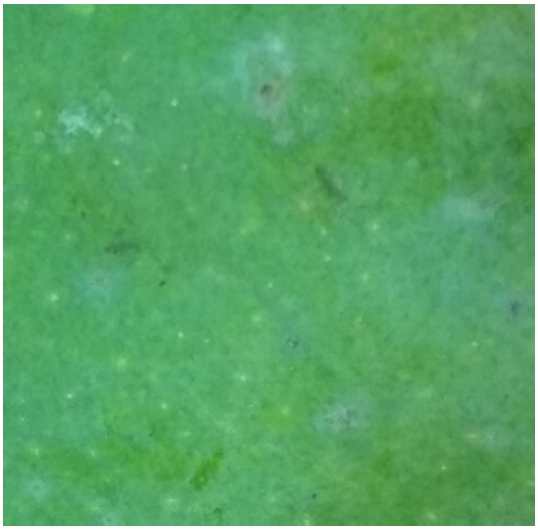

Figure 5a. Unripe mango fruit

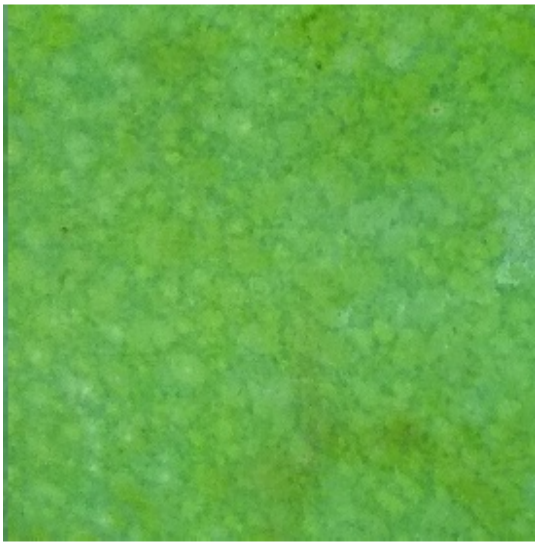

Figure 5b. Ripe mango fruit

DOI: https://doi.org/10.17509/seict.v4il.52916
p- ISSN 2774-1656 e- ISSN 2774-1699

Figure 5c. Ripe mango fruit

Figure 5d. overripe mango fruit

The first step after obtaining the mango image is feature extraction from the mango fruit. In this research, we used the following features: Red Value, standard deviation, skewness, entropy, and kurtosis as inputs. The obtained values are then normalized and grouped into 4 ripeness levels with the codes 1 (unripe), 2 (medium-ripe), 3 (ripe), and 4 (overripe). The primary parameter used to determine the ripeness level of the mango is the Red value (RGB

DOI: https://doi.org/10.17509/seict.v4il.52916
p- ISSN 2774-1656 e- ISSN 2774-1699

representation). This parameter can be logically classified and facilitates the learning process of the system.

To achieve an accurate system, several experiments need to be conducted with different percentages of training and testing data. This research performed 3 experimental scenarios with different percentages of training and testing data: 80:20, 60:40, and 50:50. These scenarios were based on previous studies (Karo et al., 2022) that yielded the best model with those compositions. Each data processing experiment was tested 10 times. The expected prediction results should be 1, 1, 2, 2, 2, 3, 3, 3, 4, and 4 in sequential order. The neuron architecture design parameters used were 2-2-1, with a learning rate of 0.1.

The activation function for the hidden layer was sigmoid, and purelin was used for the output layer. Table 1 provides an example of the testing results conducted.

Table 1. Testing result example

Testing	MSE	Epoch	Regression	*False error*	Testing Result									
1	0,038	179	0,982	1	1	1	2	2	2	4	3	3	4	4
					T	T	T	T	T	F	T	T	T	T
2	0,056	47	0,973	1	1	1	2	2	2	4	3	3	4	4
					T	T	T	T	T	F	T	T	T	T
3	0,012	26	0,995	1	1	1	2	2	2	3	3	3	3	4
					T	T	T	T	T	T	T	T	F	T
4	0,017	43	0,992	5	2	1	1	2	3	4	4	3	4	4
					F	T	F	T	F	F	F	T	T	T
5	0,017	116	0,992	5	2	1	1	2	3	4	4	3	4	4
					F	T	F	T	F	F	F	T	T	T

DOI: https://doi.org/10.17509/seict.v4il.52916
p- ISSN 2774-1656 e- ISSN 2774-1699

6	0,019	308	0,991	3	2	1	2	2	2	4	4	3	4	4
					F	T	T	T	T	F	F	T	T	T
7	0,038	58	0,982	2	1	1	2	2	2	3	4	4	4	4
					T	T	T	T	T	T	F	F	T	T
8	0,017	107	0,992	5	2	1	1	2	3	4	4	3	4	4
					F	T	F	T	F	F	F	T	T	T
9	0,017	60	0,992	5	2	1	1	2	3	4	4	3	4	4
					F	T	F	T	F	F	F	T	T	T
10	0,017	31	0,992	0	1	1	2	2	2	3	3	3	4	4
					T	T	T	T	T	T	T	T	T	T

After simulation, in the first experiment with an 8:2 ratio of training data to test data, an accuracy rate of 72% was obtained. In the second experiment, data processing was conducted with a 6:4 ratio of training data to test data. After testing, an accuracy rate of 60.90% was achieved. In the third experiment, data processing was performed with a 5:5 ratio of training data to test data, resulting in an accuracy rate of 53.75%.

From the above testing, the data training and test data process with a 5:5 ratio could not be implemented because the percentage of accuracy in the test results was almost equal to the percentage of incorrect test results. In the data training and test data process with a 6:4 ratio, it could still be implemented as the percentage of accuracy in the test results was higher than the percentage of incorrect test results. The sample data used in the 6:4 ratio consisted of 30 samples, including 6 raw mangoes, 7 ripening mangoes, 11 mature mangoes, and 6 overripe mangoes.

The best level of accuracy was achieved with an 8:2 ratio of test data to training data, with a percentage of 72%. This is because more training data was processed in this experiment compared to the previous ones.

DOI: https://doi.org/10.17509/seict.v4il.52916
p- ISSN 2774-1656 e- ISSN 2774-1699

4. CONCLUSION

After simulation, in the first experiment with a training data to test data ration of 8:2, the accuracy rate obtained was 72%. In the second experiment, data processing was conducted with a training data to test data ratio of 6:4. After testing, an accuracy rate of 60.90% was achieved. In the third experiment, data processing was performed with a training data to test data ratio 5:5. After testing, an accuracy rate of 53.75% was obtained.

Based on the above testing, the data processing with a 5:5 ratio of training data to test data cannot be implemented because the accuracy percentage is almost the same as the percentage of incorrect test results. The data processing with a 6:4 ratio can still be implemented because the accuracy rate of the test results is higher than the percentage of incorrect test results. The sample data used in the 6:4 ratio consist of 30 data, including 6 data of raw mango, 7 data of ripe mango, 11 data of mature mango, and 6 data of overripe mango. The highest accuracy rate is achieved in the 8:2 ratio of test data to training data, with a percentage of 72%. This is because more data training was processed compared to the previous two experiments.

5. AUTHORS' NOTE

The authors declare that there is no conflict of interest regarding the publication of this article. Authors confirmed that the paper was free of plagiarism.

6. REFERENCES

Adenugraha, S. P., Arinal, V., and Mulyana, D. I. (2022). Klasifikasi kematangan buah pisang ambon menggunakan metode KNN dan PCA berdasarkan citra RGB dan HSV. *Jurnal Media Informatika Budidarma*, *6*(1), 9-17.

Alcaruhban, R., Sugiantoro, B., and Prayudi, Y. (2018). Analisis pendeteksi kecocokan objek pada citra digital dengan metode algoritma sift dan histogram color RGB. *Cyber Security dan Forensik Digital*, *1*(1), 20-27.

Chen, T., and Zhong, S. (2009). Privacy-preserving backpropagation neural network learning. *IEEE Transactions on Neural Networks, 20*(10), 1554-1564.

Himmah, E. F., Widyaningsih, M., and Maysaroh, M. (2020). Identifikasi Kematangan Buah Kelapa Sawit Berdasarkan Warna RGB Dan HSV Menggunakan Metode K-Means Clustering. *Jurnal Sains Dan Informatika*, *6*(2), 193-202.

Hossain, M. A., Rana, M. M., Uddin, M. S., and Kimura, Y. (2021). Changes in organoleptic and biochemical characteristics of mango fruits treated with calcium chloride in hot water. *Journal of Horticulture and Postharvest Research, 4*(1), 37-50.

DOI: https://doi.org/10.17509/seict.v4il.52916
p- ISSN 2774-1656 e- ISSN 2774-1699

Javanmardi, S., Ashtiani, S. H. M., Verbeek, F. J., and Martynenko, A. (2021). Computer-vision classification of corn seed varieties using deep convolutional neural networks. *Journal of Stored Products Research*, *92*, 41-55.

Karo, I. M. K., Fudzee, M. F. M., Kasim, S., and Ramli, A. A. (2022). Karonese sentiment analysis: a new dataset and preliminary result. *JOIV: International Journal on Informatics Visualization*, *6*(2-2), 523-530.

Kurniawan, S. D., and Junaidi, T. (2022). Implementasi algoritma k-nearest neighbor dengan metode hue saturation value untuk pendeteksi kematangan buah jambu. *Smart Comp: Jurnalnya Orang Pintar Komputer*, *11*(3), 541-547.

Laia, F. H., Rosnelly, R., Buulolo, K., Lase, M. C., and Naswar, A. (2023). Klasifikasi kematangan buah mangga madani berdasarkan bentuk dengan jaringan syaraf tiruan metode perceptron. *Device*, *13*(1), 14-20.

Latuconsina, R. (2021). Kualitas minyak transformator ditentukan dengan pengolahan citra digital pada nilai RGB (red, green, dan blue). *Jurnal Elektrikal dan Komputer*, *2*(1), 71-78.

Nurraharjo, E. (2012). Implementasi image statistic method pada pengolahan citra digital. *Dinamik*, *17*(1), 1-5.

Paola, J. D., and Schowengerdt, R. A. (1995). A review and analysis of backpropagation neural networks for classification of remotely-sensed multispectral imagery. *International Journal of Remote Sensing, 16*(16), 3033-3058.

Salsabila, A., Yunita, R., and Rozikin, C. (2021). Identifikasi citra jenis bunga menggunakan algoritma KNN dengan ekstraksi warna HSV dan tekstur GLCM. *Technomedia Journal*, *6*(1), 124-137.

Sanjaya, C. B., and Rosadi, M. I. (2018). Klasifikasi buah mangga berdasarkan tingkat kematangan menggunakan least-squares support vector machine. *Explore IT*, *10*(2), 1-13.

Sari, J. Y., and Purnama, I. P. N. (2018). Identifikasi tingkat kematangan buah pisang menggunakan metode ektraksi ciri statistik pada warna kulit buah. *Ultimatics: Jurnal Teknik Informatika*, *10*(2), 98-102.

Shi, F., Wang, J., Shi, J., Wu, Z., Wang, Q., Tang, Z., ... and Shen, D. (2020). Review of artificial intelligence techniques in imaging data acquisition, segmentation, and diagnosis for COVID-19. *IEEE reviews in biomedical engineering*, *14*, 4-15.

Sutisna, S. P., Waluyo, R., Aldiansyah, F., and Rahmat, M. (2020). Aplikasi pengolahan citra untuk proses sortasi buah mangga berdasarkan dimensi dan bobot. *Jurnal Ilmiah Rekayasa Pertanian dan Biosistem*, *8*(1), 12-19.

Syarifah, A., Riadi, A. A., and Susanto, A. (2022). Klasifikasi tingkat kematangan jambu bol berbasis pengolahan citra digital menggunakan metode k-nearest neighbor. *JIMP (Jurnal Informatika Merdeka Pasuruan)*, *7*(1), 27-35.

DOI: https://doi.org/10.17509/seict.v4il.52916
p- ISSN 2774-1656 e- ISSN 2774-1699

Zhang, D., and Lu, G. (2004). Review of shape representation and description techniques. *Pattern recognition*, *37*(1), 1-19.

Zeb, J., Javed, M. Y., and Qayyum, U. (2007). Low resolution single neural network based face recognition. *International Journal of Computer and Information Engineering*, *1*(4), 905-909.

DOI: https://doi.org/10.17509/seict.v4il.52916
p- ISSN 2774-1656 e- ISSN 2774-1699