INTEK Jurnal Penelitian
Vol. 10, No. 1, pp. 15-21, April 2023

Papaya Fruit Quality Classification Based on Lab Color and Texture Features Using Artificial Neural Networks (ANN)

Dyah Darma Andayani [1,a], Andi Noer Aksha Rafii [1,b], Kaerurrijal Mahdar [1,c], Winda Andrayani Ahmad [1,d], Mustamin [1,*,e]
[1] [1] Computer Engineering Study Program, Universitas Negeri Makassar
[*,a]dyahdarma@unm.ac.id, [b] andiakshaltmpt@gmail.com, [c] khairurrijal13022002@gmail.com,
[d] wndahmadd@gmail.com, [e] mustamin@unm.ac.id

Abstract - **The process of sorting the quality of papaya fruit is a postharvest problem. So far, humans still do the quality sorting process conventionally or manually. It certainly has weaknesses and limitations, which require a large workforce, and the level of human perception of the quality of papaya varies. Several studies have been carried out regarding the classification of papaya fruit quality, but these studies have accuracy that can still be improved. Therefore, in this study, it is proposed to determine the quality of papaya fruit based on LAB colour and texture features using an Artificial Neural Network (ANN) algorithm. This research consists of six stages: image acquisition, pre-processing, segmentation, morphology, feature extraction and classification. The classification process for the training stage produced the highest level of accuracy in three training scenarios; namely, two techniques have 100% accuracy and 99.58% in the third scenario. Based on the best training scenario selected, the testing process produces 98.88%, the highest accuracy rate with a misclassification error of 1.12% and 69 seconds of computing time. These results indicated that the proposed method can accurately classify papaya quality based on LAB colour and texture features.**

Keywords* - *Neural Network, LAB, Quality, Papaya, Texture

I. Introduction

Papaya (Carica Papaya L) is a native fruit that grows in Central America and is widely spread in the South Pacific and other tropical areas. Papaya is capable of thriving in both arid and humid regions, encompassing elevated terrains as well as low-lying areas, particularly within tropical climates [1]. The composition of the papaya fruit exhibits variations in its chemical content based on the stage of maturity during harvest. For instance, when the fruit is unripe and green, ripe, or nearing decay, the chemical constituents differ. [2].

Papaya fruit possesses numerous nutrients that offer significant benefits to the human body. The vitamins present in papaya fruit consist of organic compounds that play a crucial role in metabolic reactions, cell growth, and overall health maintenance. This advantageous characteristic contributes to the relatively high market value of papaya fruit [3]. Meeting the demand for papaya fruit in export markets necessitates ensuring its quality in order to compete with other countries' papaya producers who also export the fruit [4].

Based on the Central Statistics Agency (BPS), the substantial consumption of papaya fruit had a notable influence on the production value, reaching 875,112 tons annually in 2017. The production further rose to 887,591 tons in 2018, consistently increasing each subsequent year. The data from the Central Statistics Agency indicates that Indonesia is projected to produce 1.17 million tons of papaya in 2021, showcasing a growth of 14.94% compared to the previous year's 1.02 million tons [5]. According to information from the Food & Agriculture Organization (FAO), Indonesia is the 5th

DOI : http://dx.doi.org/10.31963/intek.v9i2.4246

largest papaya producer in the world after Brazil, Nigeria, India and Mexico [5].

The process of sorting the quality of papaya is one of the postharvest problems; so far, humans still do the quality sorting process conventionally or manually. This method certainly has limitations and areas for improvement. Sorting manually requires a lot of human resources, not to mention the level of human perception of quality varies, and the level of human consistency only sometimes guarantees the quality of papaya because humans can experience fatigue [6].

To overcome this problem, digital image processing techniques have been carried out for sorting the quality of papaya using the Artificial Neural Network (ANN) method [7], which has been used in many past studies; there are also studies using techniques such as Fuzzy Mamdani [8], Image Processing [9], K-Nearest Neighbor [10].

In a prior investigation aiming to assess the quality of California papaya, the implementation of the ANN (Artificial Neural Network) method yielded an accuracy rate of 72.38% [7]. Similarly, a separate study focusing on the classification of papaya quality using GLCM (Gray Level Co-occurrence Matrix) texture and employing ANN achieved the highest accuracy rate of 86.11% [7]. Likewise, a study addressing the evaluation of papaya quality utilizing the fuzzy Mamdani method demonstrated an accuracy rate of 70% [8]. Furthermore, a separate research endeavor focused on estimating the quality of papaya through the application of KNN (K-Nearest Neighbors) achieved an accuracy rate of 88.88% [10].

In the preceding research, there is still room for improvement in accurately classifying papaya fruit. The limited number of datasets employed as training and test data is a contributing factor to the relatively lower accuracy achieved in the study. Increasing the dataset size would likely lead to higher accuracy in the results.

Thus, the primary objective of this study was to classify the quality of papaya fruit based on LAB color and texture features using Artificial Neural Networks. The methodology involved several stages, including image acquisition, preprocessing, segmentation, morphology, feature extraction, and classification. Additionally, the dataset utilized in this research consisted of 300 samples, which were categorized into three classes, double the size of datasets employed in previous studies. This system design holds the potential to achieve a superior level of accuracy.

II. Research Methodology

The stages in the designed system started from image acquisition, image preprocessing, image segmentation, morphological operations, feature extraction from images, and quality classification processes using artificial neural networks. The stages in the method are shown in Figure 1.

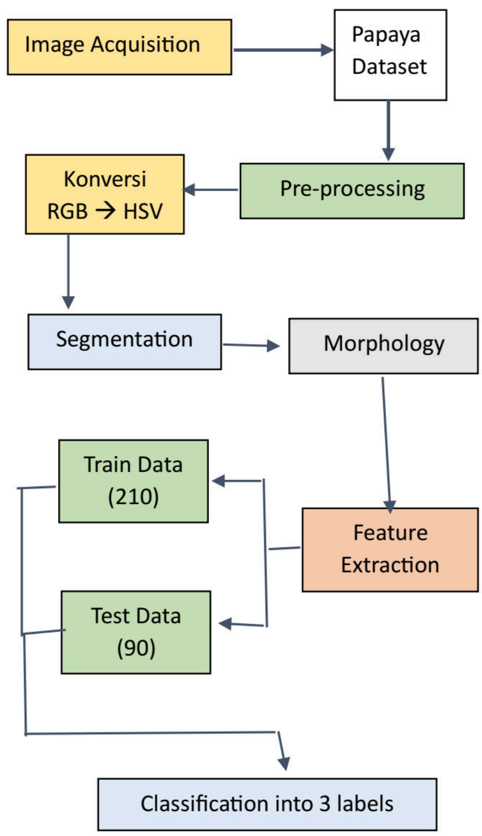


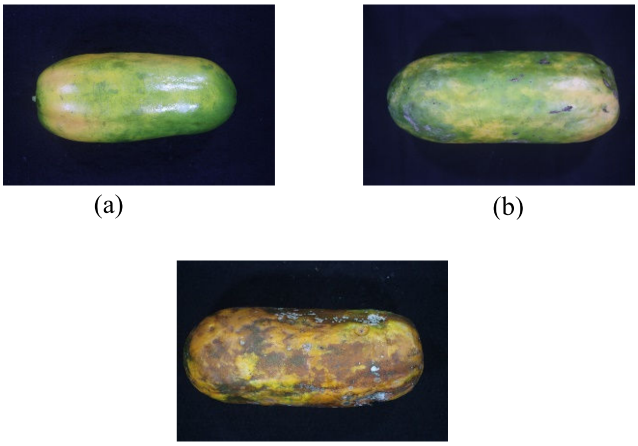
Figure 1. Block Diagram of System

A. Image Acquisition

Image acquisition was carried out on papaya fruit that had been collected, and its quality was determined into three categories, namely good, medium and bad. The process

DOI : http://dx.doi.org/10.31963/intek.v9i2.4246

of assessing the quality of papaya is carried out by local vendors who have often had direct contact with papayas so that their qualifications can be guaranteed. Papaya image data was taken using a Canon EOS 700D DSLR Camera. A total of 300 image data points were utilized in the study, comprising 100 images classified as good quality, 100 images classified as moderate quality, and 100 images classified as poor quality. The data distribution involved allocating 70% of the image data for training purposes, while the remaining 30% was designated for testing the model's performance. An example of a predetermined papaya-quality image can be seen in Figure 2.



(a) (b) (c)

Figure 2. Papaya quality (a) good, (b) moderate, (c) bad

B. Preprocessing

Preprocessing plays a crucial role in enhancing the likelihood of successful image processing [9]. As part of this process, the images were resized from their original size of 5184x3456 pixels to 1200x800 pixels, reducing the subsequent processing and computation time. Subsequently, the RGB channels within the original image were extracted to obtain the color intensity of each channel. Following this, the RGB channel values were converted to the HSV color space, which stands for Hue, Saturation, and Value.

C. Segmentation

In digital image processing, segmentation refers to the process of dividing an image into multiple regions, each with distinct attributes corresponding to the objects contained within the image [11]. Another way to understand segmentation is as the process of labeling objects and backgrounds. In this study, segmentation was conducted on images that had been converted to the HSV color space, specifically on the Hue channel.

The segmentation method employed in this research was the Otsu method, which determines a threshold value by analyzing the histogram of the image. This threshold value is then used to differentiate objects from the background, effectively segmenting the image into distinct regions [12].

D. Morphology

In digital image processing, morphological operations are often utilized to modify the appearance of objects within the original image [9]. In this study, these operations were applied to the segmented image, resulting in a morphological image. The specific morphological operations employed included dilation, closing, hole filling, and bwareaopen.

During the morphological operation, two structuring elements, or Strels, in the shape of disks were utilized. The dimensions of Strel 1 were set to 5 pixels, while Strel 2 had dimensions of 6 pixels. These structuring elements played a role in shaping and influencing the outcome of the morphological operations on the image.

E. Feature Extraction

By utilizing binary image segmentation and morphology techniques, feature extraction was performed to acquire distinct features that could effectively differentiate between different parameters related to papaya quality classes. The extracted features encompassed RGB, HSV, and LAB color features, as well as shape and texture features, which would be utilized for subsequent comparisons.

Among the extracted features, particular emphasis was placed on identifying feature values that exhibited the highest differentiating parameters for distinguishing

DOI : http://dx.doi.org/10.31963/intek.v9i2.4246

papaya quality. These selected feature values would serve as input data for the subsequent classification process. Feature extraction was carried out to obtain the feature values, and these features were stored in a database as a dataset for the training data [7]. Furthermore, the feature values would be utilized once again as input data during the training data process, enabling the determination of weights. These weights would be stored in the database as a dataset for future use during the testing stage.

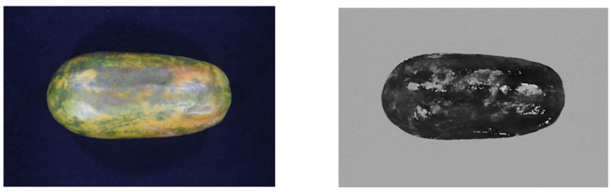
F. Classification

Following the extraction of feature values, the classification process was conducted using the Artificial Neural Network (ANN) method. This iterative method involved training and testing the data by comparing input values, obtained from the feature extraction stage, with the target values to establish a network for each image [11]. This enabled the system to determine the quality of papaya fruit based on the extracted features.

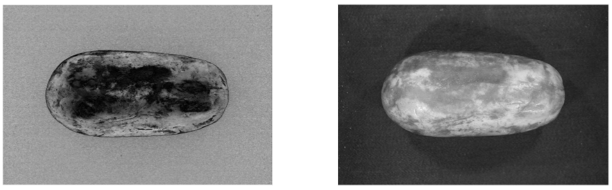
In the constructed ANN architecture, two hidden layers were utilized, with ten nodes in the first hidden layer and five nodes in the second hidden layer. Additionally, the ANN architecture employed the logistic activation function, which mapped the input values to a range between 0 and 1, facilitating the network's computations and learning processes.

III. Results and Discussion

According to Figure 2, the papaya quality classes are depicted at the image acquisition stage. The most notable distinguishing factor among the three quality classes was the condition of the fruit skin. Good quality papayas exhibited smooth and clean skin conditions, whereas medium quality papayas had a few wounds on the skin. Poor quality papayas, on the other hand, displayed extensive damage and wounds on the skin.



(a) (b)



(c) (d)

Figure 3. Channel (a) RGB, (b) hue, (c) saturation, (d) value

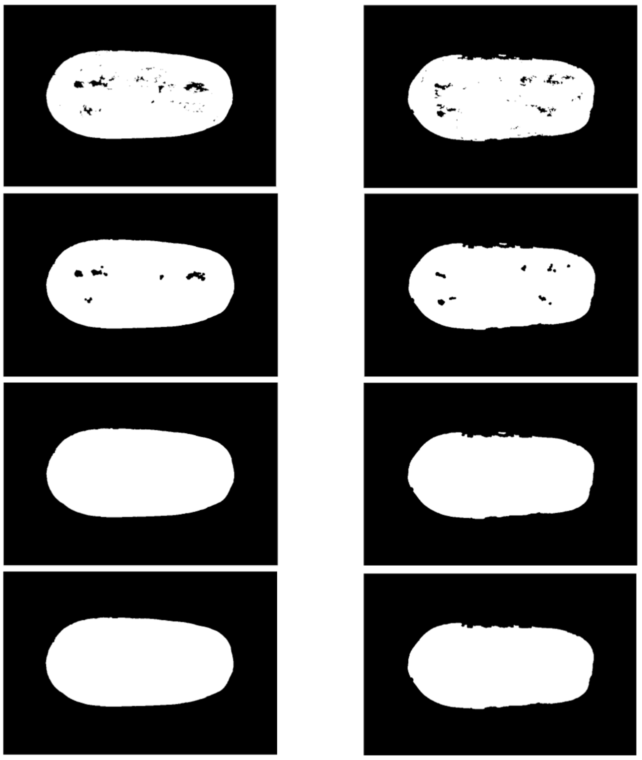
In Figure 3, the RGB channels are extracted during the preprocessing stage and subsequently converted into HSV channels. Among the three converted channels, the hue channel exhibited the most pronounced contrast between the papaya object and the background. As a result, the image derived from the hue channel was utilized for the segmentation process in order to obtain segmentation results.

(a) (b)

Figure 4. Segmentation (a) good, (b) bad

Figure 4 illustrates a comparison between good and poor segmentation results. In the case of good segmentation, a distinct boundary is observed, effectively distinguishing the objects from the background. Conversely, poor segmentation results display a slightly disorganized shape, leading to some areas of the object being undetected. These inadequate segmentation outcomes reveal the presence of empty regions within the object section, which can potentially impact the subsequent feature extraction process. To address this issue, a morphological approach becomes necessary in order to enhance the detection of objects and backgrounds, improving the overall segmentation results.

DOI : http://dx.doi.org/10.31963/intek.v9i2.4246



(a) (b)

Figure 5. Morphology (a) good (b) bad

In the morphological process shown in Figure 5, dilation, closing, hole filling, and open area operations were performed sequentially to improve object and background detection results. Based on the results of the applied morphological operations, the used morphological operations can improve the accuracy of both good and bad segmentation results.

Table 1. Training Scenario 1

Index	**Features**	**Accuracy (%)**	**ME (%)**	**Runtime (s)**
1	RGB	99.04	0.96	170
2	HSV	93.80	6.20	193
3	LAB	96.66	3.34	706
4	Shape	60.95	39.05	198
5	Texture	69.04	30.96	353

Table 1 shows the level of accuracy, misclassification error (ME) and computational time for each colour feature extraction RGB, HSV, LAB, shape features (area; perimeter; eccentricity; metric) and texture features (contrast; correlation; energy; homogeneity). Each of these feature values was used as input one by one in the Neural Network architecture.

Table 2. Training Scenario 2

Index	**Features**	**Accuracy (%)**	**ME (%)**	**Runtime (s)**
6	RGB; Shape	99.58	0.42	183
7	RGB; Texture	99.53	0.47	325
8	HSV; Shape	98.57	1.43	487
9	HSV; Texture	99.52	0.48	237
10	LAB; Shape	100	0.00	546
11	LAB; Texture	100	0.00	582

Table 3. Training Scenario 3

Index	**Features**	**Accuracy (%)**	**ME (%)**	**Runtime (s)**
12	RGB; Shape; Texture	98.09	1.91	175
13	HSV; Shape; Texture	76.66	23.34	196
14	LAB; Shape; Texture	99.04	0.96	527

Table 2 shows the accuracy and runtime of the two feature combinations used as training scenarios. Table 3 combines three different feature values to be used as training scenarios. Three combinations of all the training scenarios were selected with the highest accuracy, the lowest misclassification error (ME) and the shortest runtime: a combination of RGB features; Shape, LAB features; LAB form and features; Texture.

DOI : http://dx.doi.org/10.31963/intek.v9i2.4246

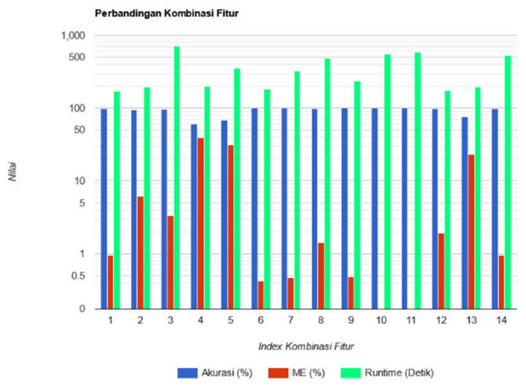


Figure 6. Comparison of Features Index Combination

Figure 6 provides a comparison of accuracy, misclassification error (ME), and runtime for different feature combination scenarios during the training process. It was observed that index 6, which combined RGB features with shape, achieved an accuracy of 99.58%, ME of 0.42%, and a runtime of 183 seconds. Indexes 10 and 11, which combined LAB features with shape and textures, respectively, resulted in 100% accuracy, 0.00% ME, and runtimes of 548 seconds and 582 seconds. These three feature combinations were identified as the most accurate in determining the quality of papaya fruit. Consequently, this feature combination was selected for use in the subsequent testing process.

Table 4. Result of Scenario Testing

Index	Features	Accuracy (%)	ME (%)	Runtime (s)
6	RGB; Shape	96.66	3.34	33
10	LAB; Shape	96.66	3.34	66
11	LAB; Texture	98.88	1.12	69

Table 4 presents the testing process using the three chosen feature combinations from the training process. Among these combinations, the LAB feature combination with textures was deemed to possess the highest level of accuracy and the lowest misclassification error (ME), while its runtime was comparable to the other feature combinations. This evaluation of accuracy, ME, and runtime led to the determination that the LAB feature combination with textures was the most favorable choice among the three combinations.

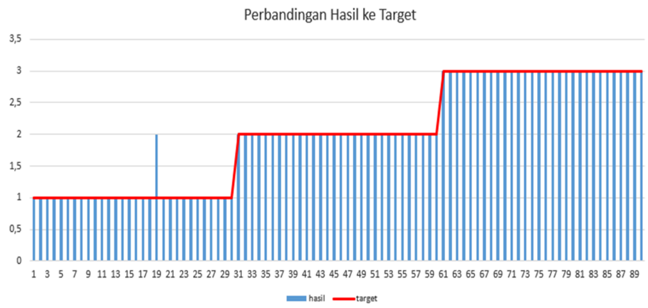


Figure 7. Comparison of Result to Target

Figure 4 displays a comparison between the results of quality classification and the actual quality classes. It reveals that there was an instance where an image of papaya was misclassified. The image, which should have been categorized as quality target 1, was mistakenly identified by the system as belonging to quality class 2. This misclassification contributed to the overall accuracy of 98.88%. Considering the size of the papaya image dataset used, this level of accuracy is considered remarkably high.

IV. Conclusion

The results of the papaya quality classification using an artificial neural network algorithm indicated high accuracy rates. The training process achieved a perfect accuracy of 100% when utilizing a combination of LAB features with shape and LAB features with texture. Additionally, a high accuracy of 99.58% was obtained with various combinations of RGB features with shape. During the testing process, the selected features, specifically LAB features with texture, yielded an accuracy of 98.88% and a misclassification error (ME) of 1.12%. The computation time for this process was 69 seconds.

Further research is recommended to explore the utilization of additional features in determining the quality of papaya fruit. Additionally, it is advisable to

DOI : http://dx.doi.org/10.31963/intek.v9i2.4246

expand the dataset used, ensuring it is either the same size or larger than the dataset employed in this study. These measures may contribute to enhancing the accuracy and reliability of future papaya quality classification studies.

References

[1] N. J. H. Smith, J. T. Williams, D. L. Plucknett, and J. P. Talbot, *Tropical forests and their crops*. Cornell University Press, 2018.

[2] H. Kaur, B. K. Sawhney, and S. K. Jawandha, "Evaluation of plum fruit maturity by image processing techniques," *J Food Sci Technol*, vol. 55, pp. 3008–3015, 2018.

[3] I. Salinas, J. J. Hueso, and J. Cuevas, "Active control of greenhouse climate enhances papaya growth and yield at an affordable cost," *Agronomy*, vol. 11, no. 2, p. 378, 2021.

[4] N. R. NMM, C. Suntharalingam, and M. F. Othman, "Competitiveness of Malaysia's fruits in the global market: Revealed comparative advantage analysis," *Malaysian Journal of Mathematical Sciences*, vol. 11, pp. 143–157, 2017.

[5] S. Sadya, "Produksi Pepaya di Indonesia Capai 1,17 Juta Ton pada 2021," 2022. https://dataindonesia.id/sektor-riil/detail/produksi-pepaya-di-indonesia-capai-117-juta-ton-pada-2021 (accessed Dec. 03, 2022).

[6] K. K. Patel, *Studies on Destructive and Non-Destructive Quality Evaluations of Mango Fruit*. Notion Press, 2020.

[7] J. Naranjo-Torres, M. Mora, R. Hernández-García, R. J. Barrientos, C. Fredes, and A. Valenzuela, "A review of convolutional neural network applied to fruit image processing," *Applied Sciences*, vol. 10, no. 10, p. 3443, 2020.

[8] M. E. Al Rivan and J. Suherman, "Penentuan Mutu Buah Pepaya California (Carica Papaya L.) Menggunakan Fuzzy Mamdani," *ELKHA: Jurnal Teknik Elektro*, vol. 12, no. 2, pp. 76–83, 2020.

[9] W. D. Cahyono and D. Wirawan, "KAJIAN SIFAT FISIK BUAH PEPAYA (Carica papaya L.) MENGGUNAKAN PENGOLAHAN CITRA".

[10] F. Wibowo, D. K. Hakim, and S. Sugiyanto, "Pendugaan Kelas Mutu Buah Pepaya Berdasarkan Ciri Tekstur Glcm Menggunakan Algoritma K-Nearest Neighbors," *Jurnal Nasional Pendidikan Teknik Informatika: JANAPATI*, vol. 7, no. 1, pp. 100–106, 2018.

[11] A. Firlansyah, A. B. Kaswar, and A. A. N. Risal, "Klasifikasi Tingkat Kematangan Buah Pepaya Berdasarkan Fitur Warna Menggunakan JST," *Techno Xplore: Jurnal Ilmu Komputer dan Teknologi Informasi*, vol. 6, no. 2, pp. 55–60, 2021.

[12] S. S. Chouhan, A. Kaul, and U. P. Singh, "Image segmentation using computational intelligence techniques," *Archives of Computational Methods in Engineering*, vol. 26, pp. 533–596, 2019.

DOI : http://dx.doi.org/10.31963/intek.v9i2.4246