Nondestructive Banana Ripeness Classification using Neural Network

Maimunah
Department of Informatics Engineering
Universitas Muhammadiyah Magelang
Magelang, Indonesia 56172
maimunah@ummgl.ac.id

Rahmadya Trias Handayanto
Computer Engineering
Universitas Islam 45
Bekasi, Indonesia 17113
rahmadya.trias@gmail.com

Herlawati
Informatics Engineering
Universitas Bhayangkara Jakarta Raya
Bekasi, Indonesia 17121
mrs.herlawati@gmail.com

***Abstract*—Banana contributes to the total fruit production in Indonesia. To ensure the quality of a banana, Indonesian government with its national standard agency (BSN) has issued the national standard of fruit. This study proposed the classification method of a banana using the neural network based on color and texture. Data was collected before model creation with numbers of banana of ripeness level 2, 3, and 4 were 30, 30, and 45 respectively. Three neural network model with three different variables, i.e. color image and texture. Testing result show a high accuracy result of 95.24%.**

Keywords—ambon banana, ripeness, classification, color index.

I. Introduction

Banana, a fruit which contributes to total national fruit production, supplies not only national but also international market. Therefore, a quality control for this fruit is needed in order to gain a competitive product. Government, through National Standard Bureau, has issue a standard for banana quality, SNI 7422:2009.

Various types of banana are available in Indonesia. One of the famous types is ambon banana. This type usually appears in a food court with other foods. Yellow ambon banana when ripe has a yellow color with beige or white inside with sweet taste and strong smell. Ambon banana can also be used as a first food for a baby. This fruit has 15-25 kg weight and contains 10 to 14 pieces. It is a long banana (15-20 cm) with 3.45 cm diameter. Ambon banana can be used for other foods, e.g. food drink, jam, banana flour, smoked banana, etc. [1].

Banana (Musa Paradisiaca) is a kind of fruit that full of vitamin, mineral, and carbohydrate. In Indonesia, banana is usually planted in the garden without any special treatment. Therefore, the quality is low and cannot have competitive advantages for export. Banana was cultivated in 80 to 100 days according to its variety. There are two kinds method of cultivation: i) based on banana heart cut and ii) counting the plant's age or by seeing directly the fruit. Before cultivation, it is important to know the age of the banana tree as a main quality factor. If the banana age too old when cultivation, it has bad storability even the taste is sweet and fragrant. In the other hand, if the age is not too old, the quality is low (taste and smell). Therefore, age level when cultivating the banana should be related to marketing range and the purpose of banana use.

II. Related Research

Banana classification can be used by destructive and nondestructive analysis. Destructive method uses some chemical substrates of a banana sample. The results are information about total solution density, starch content, and acid content. But this method need samples and facilities [1].

Nondestructive method in classifying ripeness level by seeing the skin color, texture, and other external information of a banana [2]–[6]. Another study used smell from Banana with fuzzy C-means and neural network [7]. Indonesian standard number: 7422:2009 for banana ripeness has been widely used [8].

In neural networks, weight values are obtained from a particular learning algorithm, whereas in Naïve Bayes algorithm the probability calculation is used with a simplification, i.e. assumption that all attribute values are conditionally independent for all output values given. The neural network approach is more refined, more mathematical and potentially far more accurate and reliable in accomplishing this task [9]

Classification of the maturity stage of ambon banana has been carried out using Naïve Bayes with an accuracy of 90.48% in previous study [10]. With the same data, in this study use the nondestructive banana ripeness classification was proposed through neural network classifier.

III. Methodology

Fig 1 shows the research framework with some steps, i.e. image acquisition, feature extraction, neural network modeling, neural network training, ripeness classification.

A. Image Acquisition

The research used 105 Ambon Banana image. Ambon bananas that are picked in100 days after blooming (HSBM) are fruit with the first stage of maturity. After picking and when doing the separation of bananas with bunches, the banana sap did not touch the fruit skin.

This study captured 15 units of Banana. The image was taken when the banana entered the second stage, which was six days after the first maturity stage, the image is taken at night around 8.00 to 9.00 pm.

The second stage of maturity imagery was taken on days 6 and 7 of the HSBM at night. Image acquisition was done on days 8 to 9 of HSBM to get the third stage of maturity. Fourth stage of maturity retrieved the image on the 10th day to the 12th of HSBM at night (Table 1).

Image acquisition is the initial stage to obtain digital images. The images were captured using a pocket type digital camera with a pixel quality of 16.1 Mega Pixels and an LCD size of 3 inches with a weight of 135 grams. A mini studio was used capturing the banana image.

A mini studio consists of two essential types of 5 watts with the length of 40 cm, 30 cm width, and a height of 35 cm. This arrangement made the studio can capture with the same lighting from two existing lights, the distance of the camera to the same object is 35 cm, and the same type of camera.

Image acquisition was done using a pocket type digital camera. More than 105 banana images were captured consisted of 3 classes maturity based on the ministry of agriculture regulation standard (SNI 7422: 2009). Maturity stage 2, stage 3, and stage 4 have samples of 30, 30, and 45 images, respectively.

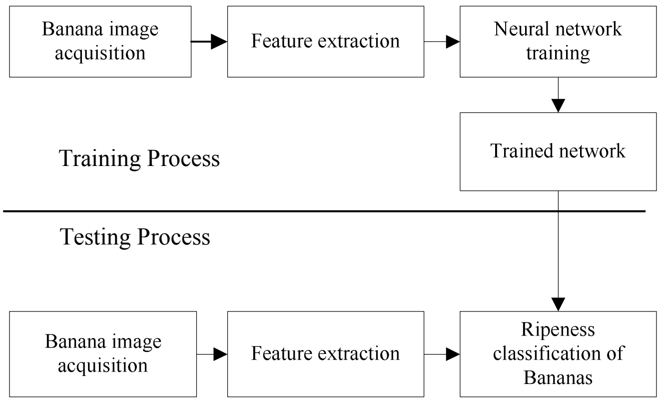

Fig. 1. Research Framework

B. Feature Extraction

Image was processed using Matlab R2014b to get the RGB value which is the average of all pixels. Then the value was normalized (rgb) by dividing each value by the number 255. The rgb value was then converted to HSV L a * b * and a co-occurrence matrix with 0 degree orientation angle was used to obtain the value of entropy, contrast, energy and homogeneity. To get the values of r, g, b which were then converted to HSV values can be obtained by the following equation:

$$r = \frac{R}{R+G+B} \quad (1)$$

$$g = \frac{G}{R+G+B} \quad (2)$$

$$b = \frac{B}{R+G+B} \quad (3)$$

Furthermore, to get the HSV value, RGB value was converted with the following equation:

$$V = max \quad (4)$$

$$S = \begin{cases} 0, if\ max = min \\ \frac{max-min}{V}, otherwise \end{cases} \quad (5)$$

$$H = 0\ if\ max = min \quad (6)$$

$$H = \begin{cases} 60^0 \times \left(\frac{G-B}{max-min}\right), if\ max = R \\ 60^0 \times \left(\frac{B-R}{max-min} + 2\right), if\ max = G \\ 60^0 \times \left(\frac{R-G}{max-min} + 4\right), if\ max = B \end{cases} \quad (7)$$

$$if\ H < 0\ then\ H = H + 360 \quad (8)$$

Where $max = max(R, G, B)$ and $min = min(R, G, B)$.

When the RGB and HSV values have been, the next step was to convert the RGB value to get the L * a * b value. The value of x was R'G 'or B' value. A function $f(x)$ indicated the conversion value of sR, sG and sB. The sRGB value was then converted to the CIE XYZ color model using the equation:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.4124 & 0.3756 & 0.1805 \\ 0.2126 & 0.7152 & 0.0722 \\ 0.0193 & 0.1192 & 0.9505 \end{bmatrix} \begin{bmatrix} sR \\ sG \\ sB \end{bmatrix} \quad (9)$$

To calculate the value of L * a * b from CIEXYZ, the following equation was used:

$$L = 116 * f\left(\frac{Y}{Y_n} - 16\right) \quad (10)$$

$$a^* = 500 * \left(f\left(\frac{X}{X_n}\right) - f\left(\frac{Y}{Y_n}\right)\right) \quad (11)$$

$$b^* = 200 * \left(f\left(\frac{Y}{Y_n}\right) - f\left(\frac{Z}{Z_n}\right)\right) \quad (12)$$

When all values of RGB, HSV and L * a * b have been obtained, then the value of texture features was calculated.

The Entropy, contrast, energy, and homogeneity terms can be explained as follows. Entropy is used to measure the irregularity of an object, contrast is used to measure differences between points of an object, energy is used to measure the texture diversity of an object, and homogeneity is used to measure the uniformity of an object.

The texture measurement component can be taken using the equation:

$$energy = \sum_{i,j=0}^{N-1} P_{i,j}^2 \quad (13)$$

$$entropy = \sum_{i,j=0}^{N-1} P_{i,j}(-\ln P_{i,j}) \quad (14)$$

$$contrast = \sum_{i,j=0}^{N-1} P_{i,j}(i-j)^2 \quad (15)$$

$$homogeneity = \sum_{i,j=0}^{N-1} \frac{P_{i,j}}{1+(i-j)^2} \quad (16)$$

Where i and j represent pixel column and row column respectively.

Thirteen variables, namely r, g, b, h, s, v, L, * a, b *, and texture features in the form of homogeneity, energy, contrast and entropy, were used as predictors of the maturity of ambon bananas. Estimator variables that are used as input values are made in 3 randomly selected input models with the aim of getting the highest accuracy that can be seen in Table 1.

Furthermore, the data was divided into two groups separately, i.e. training and testing data. Training data of 80% in each stage of maturity, amounting to 84 data and used for testing data by 20% in each stage of maturity, which amounted to 21 data.

Each stage of maturity was labeled, i.e. class 1 for stage of maturity 2, class 2 for stage of maturity 3, and class 3 for stage of maturity 4.

TABLE I. INPUT MODEL

Variables	Model 1	Model 2	Model 3
R	√	√	√
G	√	√	√
B	√	√	√
H	√		
S	√		
V	√	√	
L	√		
a*	√	√	√
b*	√	√	
Entropy	√	√	√
Homogeneity	√	√	√
Energy	√	√	
Contrast	√		

C. Classification

Artificial Neural Network (ANN) or neural network (NN) is a computing system where architecture and operations are inspired by knowledge of biological nerve cells in the brain, so that they are able to carry out certain tasks, especially pattern recognition with high effectiveness [11]. NN resembles the human brain, knowledge is obtained by the network through the learning process and the strength of the relationship between nerve cells (neurons) known as synaptic weights used to store knowledge. NN can be described as a mathematical and computational model for non-linear approximation functions, data classification, clusters and non-parametric regression or as a simulation of a collection of biological neural models. The neural model is demonstrated by its ability in emulation, analysis, prediction and association. Many decision support system (DSS) have used NN for their system [12].

IV. RESULT AND DISCUSSION

Sampling Image is a method for obtaining input values in the form of RGB, HSV, LAB and texture features of energy, entropy, homogeneity and contrast that will be processed using the Naïve Bayes method to determine the maturity stage of ambon banana. Image samples acquisition was done through MATLAB software.

An image sample (Fig 2) was processed using MATLAB R20014b to get an RGB value which was the average of all pixels. Then the value was normalized to RGB by dividing each value by the number 255. The RGB value was then converted into HSV and CIEXYZ colour space before finally the value in the CIEXYZ color space is converted to CIELa * b *. To get the value of the texture features using the GLCM method with an orientation angle of 0^{o} using equation 13-16.

Input values used for the classification of ambon banana maturity stages are 13 estimating variables, namely, r, g, b, h, s, v, L, * a, b *, and texture features in the form of homogeneity, energy, contrast and entropy. The estimator variable used consists of 3 input models whose variables were randomly selected in order to find the highest accuracy.

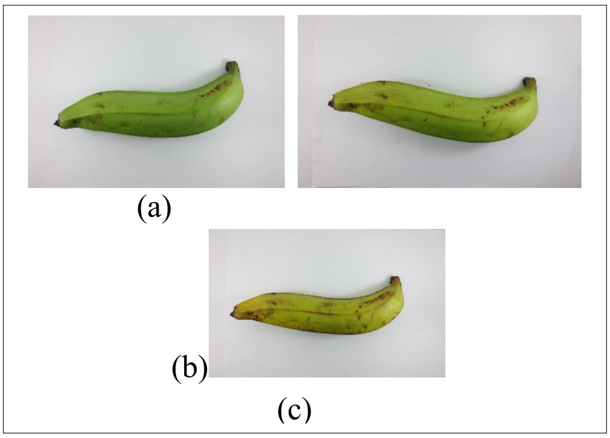

Fig. 2. Maturity Level 2 (a), Level 3 (b), and Level 4 (c)

The ambon banana image sample data was extracted using the Matlab application to obtain estimator values. Then the data is divided into two groups of data that are foreign to each other, namely training data and testing data. The training data was taken at 80% at each stage of maturity, the total training data amounted to 84 Ambon banana image data. The training phase was the stage to conduct training of training data to build an output category of ambon banana maturity stage. Testing data taken by 20% at each stage of maturity, the total training data amounted to 21 ambon banana image data.

In the training phase requires input and targets to get the most optimal weights. The training parameters specified are listed in Table II.

TABLE II. TRAINING PARAMETERS

Parameters	Parameter Value
Maximum *Epoch*	1000 Epoch
Learning Rate	0.01
MSE	0.0001
Activation Function	tansig, purelin, trainlm

Based on confusion matrix (Table III). Neural networks model was able to classify bananas based on their maturity stage with an accuracy of 95.24% for all model parameters used in 0:00:02 second. This result is better than using Naïve Bayes (with accuracy of 90.48%) for all three models [10].

TABLE III. CONFUSION MATRIX

		Prediction Class		
		Level 2	Level 3	Level 4
Actual Class	Level 2	6	0	1
	Level 3	0	5	1
	Level 4	0	0	9

V. Conclusions

Some soft computing methods have been widely used for classification. This study showed the ability of neural networks in maturity level classification. Result indicated that the neural networks had better performance compared to Naïve Bayes. For the future work, the similar model for other fruits will be conducted.

Acknowledgment

The authors thank to Universitas Muhammadiyah Magelang, Universitas Islam 45 Bekasi, and Universitas Bhayangkara Jakarta Raya. The authors also thanks to reviewers for their insightful comments.

References

[1] S. Prabawati, Suyanti, and D. A. Setyabudi, "Teknolog? Pascapanen dan Tekn?k Pengolahan Buah Pisang," *Balai Besar Penelit. dan Pengemb. Pascapanen Pertan.*, vol. 49, pp. 69–73, 2008.

[2] Indarto and Murinto, "Deteksi Kematangan Buah Pisang Berdasarkan Fitur Warna Citra Kulit Pisang Menggunakan Metode Transformasi Ruang Warna HIS (Banana Fruit Detection Based on Banana Skin Image Features Using HSI Color Space Transformation Method)," *J. Ilm. Inform.*, vol. V, pp. 15–21, 2017.

[3] S. E. Adebayo, N. Hashim, K. Abdan, M. Hanafi, and M. Zude-Sasse, "Banana Quality Attribute Prediction and Ripeness Classification Using Support Vector Machine," *ETP Int. J. Food Eng.*, vol. 3, no. 1, pp. 42–47, 2017.

[4] E. Llobet, E. L. Hines, J. W. Gardner, and S. Franco, "Non-destructive banana ripeness determination using a neural network-based electronic nose," *Meas. Sci. Technol.*, vol. 10, no. 6, pp. 538–548, 1999.

[5] H. Vaviya, V. Vishwakarma, A. Yadav, and N. Shah, "Identification of Artificially Ripened Fruits Using Machine Learning," *SSRN Electron. J.*, 2019.

[6] J. C. Hou, Y. H. Hu, L. X. Hou, K. Q. Guo, and T. Satake, "Classification of ripening stages of bananas based on support vector machine," *Int. J. Agric. Biol. Eng.*, vol. 8, no. 6, pp. 99–103, 2015.

[7] F. M. A. Mazen and A. A. Nashat, "Ripeness Classification of Bananas Using an Artificial Neural Network," *Arab. J. Sci. Eng.*, 2019.

[8] Y. P. Wiharja and A. Harjoko, "IJEIS: Indonesian journal of electronics and instrumentation systems.," *IJEIS (Indonesian J. Electron. Instrum. Syst.*, vol. 4, no. 1, pp. 57–68, 2014.

[9] A. Kumar Sharma, S. Kumar Prajapat Assistant Professor, and M. Aslam, "A Comparative Study between Naïve Bayes and Neural Network (MLP) Classifier for Spam Email Detection," *Int. J. Comput. Appl.*, pp. 975–8887, 2014.

[10] D. Yulianto, R. N. Whidhiasih, and M. Maimunah, "Klasifikasi Tahap Kematangan Pisang Ambon Berdasarkan Warna Menggunakan Naive Bayes," *PIKSEL Penelit. Ilmu Komput. Sist. Embed. Log.*, vol. 5, no. 2, pp. 60–67, 2018.

[11] L. Fausett, *Fundamentals of Neural Networks*, no. 1. 1994.

[12] S. Haykin, "Neural networks: a comprehensive foundation by Simon Haykin, Macmillan, 1994, ISBN 0-02-352781-7.," *The Knowledge Engineering Review*, vol. 13, no. 4. pp. 409–412, 1999.