

Comparison of Neural Network Methods for Classification of Banana Varieties (Musa Paradisiaca)

Zilvanhisna Emka Fitri¹, Wildan Bakti Nugroho¹, Abdul Madjid² and Arizal Mujibtamala Nanda Imron³

¹Teknik Informatika, Jurusan Teknologi Informasi, Politeknik Negeri Jember

²Budidaya Tanaman Perkebunan, Jurusan Produksi Pertanian, Politeknik Negeri Jember
Jl. Mastrip PO.BOX 164 Jember, 68121

³Teknik Elektro, Jurusan Teknik Elektro, Universitas Jember
Jl. Kalimantan No. 37, Kampus Tegalboto, Jember, 68121
e-mail: zilvanhisnaef@polije.ac.id

Abstract—Every region in Indonesia has a massive diversity of banana species, but no system records information about the characteristics of banana varieties. This research aims to make an encyclopedia of banana types that can be used for learning by classifying banana varieties using banana images. This banana variety classification system uses image processing techniques and artificial neural network methods as classification methods. The varieties of bananas used are Merah bananas, Mas Kirana bananas, Klutuk bananas, Raja bananas and Cavendis bananas. The parameters used are color features (Red, Green, and Blue) and shape features (area, perimeter, diameter, and fruit length). The intelligent system used is the backpropagation method and the radial basis function neural network. The results showed that both approaches could classify banana varieties with an accuracy rate of 98% for Backpropagation and 100% for the radial basis function neural network.

Keywords: *banana varieties, digital image processing, backpropagation, radial basis function neural network*

Abstrak—Setiap daerah di Indonesia memiliki keanekaragaman jenis pisang yang sangat banyak, namun belum ada sistem yang mencatat informasi tentang ciri-ciri varietas pisang. Tujuan dari penelitian ini adalah membuat ensiklopedia jenis pisang yang dapat digunakan untuk pembelajaran dengan cara mengklasifikasikan varietas pisang menggunakan citra pisang. Sistem klasifikasi varietas pisang ini menggunakan teknik pengolahan citra dan metode jaringan saraf tiruan sebagai metode klasifikasi. Jenis pisang yang digunakan adalah pisang merah, pisang Mas Kirana, pisang Klutuk, pisang Raja dan pisang Cavendis. Parameter yang digunakan adalah fitur warna (*Red, Green dan Blue*) dan fitur bentuk (luas area, perimeter, diameter dan panjang buah). Sistem cerdas yang digunakan adalah metode *backpropagation* dan *radial basis function neural network*. Hasil penelitian menunjukkan bahwa kedua metode tersebut mampu mengklasifikasikan varietas pisang dengan tingkat akurasi sebesar 98% untuk *Backpropagation* dan 100% untuk *radial basis function neural network*.

Kata kunci: *varietas pisang, pengolahan citra digital, backpropagation, radial basis function neural network*

I. INTRODUCTION

Banana (*Musa spp.*) is an herbaceous fruit plant that is easy to find, there are many varieties and the price is very affordable and comes from regions in Southeast Asia including Indonesia [1]. Bananas are divided into four groups based on the types of varieties and their uses, namely: 1) bananas that are eaten without cooking, namely *M. paradisiaca var sapientum*, *M. nana* or also known as *M. cavendishii*, *M. sinensis*, for example ambon bananas, susu bananas, raja bananas, cavendis bananas, barangan bananas and mas bananas; 2) bananas that are eaten after the fruit is cooked, namely *M. paradisiaca forma typica* or also known as *M. paradisiaca normalis*, for example

nangka bananas, tanduk bananas and kapok bananas; 3) banana seeds, namely *M. brachycarpa* which in Indonesia is used for its leaves, for example batu bananas and klutuk bananas; 4) bananas are taken for fiber, for example manila bananas [2]. According to *Badan Pusat Statistik* (BPS) data, there has been an increase in banana production in East Java Province from 2015 to 2019. Total production in 2015 was 1,629,437 tons and total production increased by 2,116,974 tons in 2019, with a percentage growth of 2.77% in 2018-2019 [3].

Every region in Indonesia has many varieties of bananas, giving rise to diversity of banana varieties. However, there is no system that records information regarding the characteristics of banana varieties. The purpose of this

research is to make an encyclopedia about the kinds of banana varieties that can be used for learning about the types of bananas through their pictures. This encyclopedia can classify types of bananas through photos of banana fruit, it is hoped that people who do not understand the types of bananas can be helped. Besides bringing up the classification of banana types, there is also a brief explanation of their characteristics, both in terms of taste, size and others.

Research related to bananas has been carried out, the topics that are mostly researched are identification of the level of maturity and identification of the type of banana. The focus of this research is on the classification of types or varieties of bananas, so here are some references that we use in this study. First, Ten classes of bananas were classified, namely Mas bananas, Barangan bananas, Merah bananas, Susu bananas, Tanduk bananas, Kepok bananas, Batu bananas, Awak bananas, Raja bananas and Ambon bananas with a system accuracy rate of 80% using the K - Nearest Neighbor (KNN) method [4]. Second, the KNN method is still used in this study but the HSV color parameter in classifying five types of bananas (Ijo bananas, Sobo Pipit bananas, Tandes bananas, Raja Uli bananas and Raja bananas) with an accuracy rate of 82% [5]. Third, the identification system for Muli bananas, Ambon bananas, and Kepok bananas with features used, namely weight, volume, area, roundness, Red (R), Green (G), Blue (B) and diameter with a system accuracy level of 100% using a neural network model (backpropagation) [6]. Fourth, the backpropagation method is also used to identify Lilin bananas, Mas Kirana bananas and Raja Sereh bananas with an accuracy level of 100% of the training results, and testing of 80% of training data of 50 banana images and test data of 30 banana images .

Based on the above references, the backpropagation method is able to classify banana varieties better than the K-nearest neighbor method so that the researchers use the backpropagation method to identify other banana varieties such as Kirana Mas bananas, Raja bananas, Klutuk bananas, red bananas and Cavendis bananas. In this research using a different image processing techniques with the previous research [6], [7] on segmentation and feature extraction process used. This research also compares the effectiveness of the backpropagation method with other artificial neural network methods such as the Radial Basis Neural Network (RBFNN) for classifying banana varieties. The RBFNN method was chosen because RBFNN is able to classify welding defect images with an accuracy rate of 91.67% [8] and able to classify shallots and onions with an accuracy rate of 100% on 100 test data images[9].

II. METHODS

In this study, four varieties of bananas were used, namely Merah bananas, Mas Kirana bananas, Klutuk bananas, Raja bananas and Cavendis bananas. There are six stages in classifying varieties of bananas. First, the banana image sample is carried out a processing process,

namely the cropping process which functions to speed up the computation process. Second, the process of solving RGB colors is carried out, this is because the RGB color image is difficult to segment [10], [11]. The RGB color space consists of three color components, namely red, green and blue components, each of these components has a value range of 8 bits so that the RGB color space has a value of 24 bits and this value is very large [12] and of course makes the computational load high, so that the image with the RGB color space should be converted to another color space before the segmentation process is carried out. After the preprocessing stage, the stages are divided into two, namely the extraction of the RGB color features and the color conversion process to the HSV color space (third stage). The fourth stage is after the conversion process of the color to the HSV color space is carried out, then the image segmentation process is carried out which aims to separate the object and the background. Fifth, we get the parameters which will be carried out by the morphological feature extraction. Sixth, a classification process is carried out using the backpropagation method to classify banana varieties into five classes based on color and morphological features, this stage is shown in Figure 1.

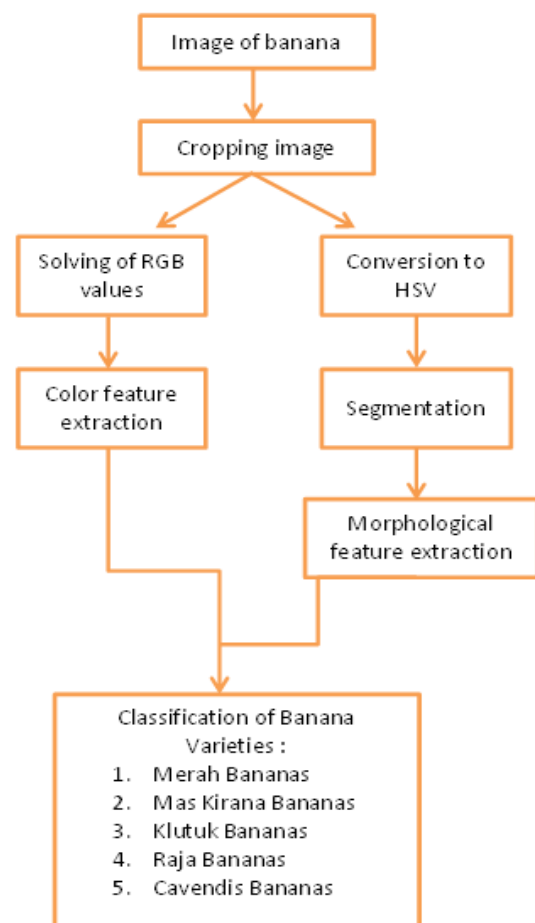


Figure 1. The block diagram of the classification system

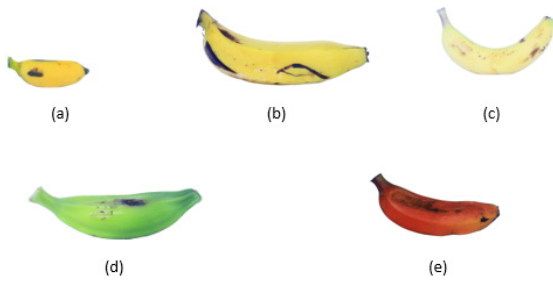


Figure 2. Banana varieties (a) Mas Kirana bananas, (b) Raja bananas, (c) Cavendis bananas, (d) Klutuk bananas, and (e) Merah bananas

A. Sample image of banana

In the initial stage, the researchers bought bananas from banana sellers in the fruit market and were randomly selected without considering the size of the bananas. Then the image is taken using a mini studio box and a DSLR camera with a resolution of 18 MP so that the data we use in this study is the researchers' private dataset. The data used consisted of five classes, Merah bananas, Mas Kirana bananas, Klutuk bananas, Raja bananas and Cavendis bananas as shown in Figure 2.

B. Image cropping

Banana sample image with initial size of 5184 x 3456 pixels. The next step is to crop the image to a smaller 612 x 408 pixels. The cropping process aims to reduce the computational burden.

C. Solving of RGB values

The next stage is the process of solving the RGB components, aiming to find the RGB components that best represent the object under study. The process of solving the RGB components is shown in Figure 3

D. Conversion to the HSV color space

The RGB color space is difficult to segment so that the color conversion process is carried out to another color space, for example HSV. In addition, the HSV color space is a color space whose model is close to that of human vision.

E. Segmentation

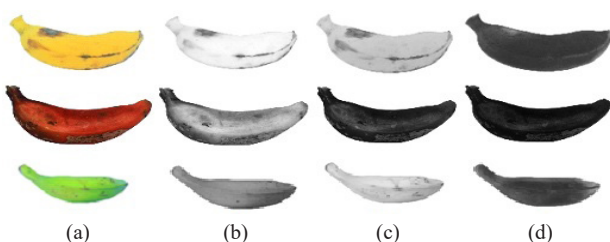


Figure 3. (a) Original image, (b) Red component image, (c) Green component image and (d) Blue component image

The main purpose of the segmentation stage is to separate the object from its background [13], this process is also often called the thresholding process. This process produces a binary image that has two values, “0” for the black region and “1” for the white region as follows:

$$S(x,y) = \begin{cases} 0 & \text{if } f(x,y) < T \\ 1 & \text{if } f(x,y) \geq T \end{cases} \quad (1)$$

where $f(x,y)$ is the gray value at a pixel coordinate, T is the threshold value determined from the histogram analysis and $S(x,y)$ is the binary value of the segments result at a pixel coordinate.

F. Feature extraction

Feature extraction used in this study consists of color feature extraction and morphological feature extraction. This is because the obtained two characteristics that can distinguish five varieties of bananas.

G. Classification

Backpropagation is an artificial neural network method which is divided into two stages, namely the training stage and the testing stage. backpropagation training methods including supervised learning [13]. Supervised learning is a learning model in which the data already has a target, if the classification results do not match the predetermined target, the weight will be updated in each hidden layer. Updates the weights are expected to improve the accuracy of the system [11], [14]. However, apart from using the backpropagation method, the researcher also made comparisons with other neural network methods, for example the Radial Basis Function Neural Network (RBFNN).

The fundamental difference between the two methods lies in the activation function of the hidden layer. MLPNN uses the sigmoid activation function while RBFNN uses Gaussian activation function as follows:

$$h(x) = \varnothing \left((x-c)^t R^{-1} (x-c) \right) \quad (2)$$

III. RESULTS AND DISCUSSION

In this research, image processing techniques specifically the segmentation process plays an important role the successful classification of varieties of bananas. Before performing the segmentation process, a color conversion process is carried out from the RGB color space to the HSV color space. The HSV color space has a model similar to that of human vision. After the HSV color conversion process is carried out, then the HSV component is resolved to determine the component image that best represents the research object as shown in Figure 4.

Figure 4 shows that the image that best represents banana objects is the saturation component image where the object and the background are clearly visible in color

difference, while the hue component image of the object tends to be darker towards the background. This also applies to the value component image; the object has a value that is close to the background or value 1 (white).

The next process is the segmentation process which converts the saturation component image into a binary image using the equation formula 1. The threshold value (T) is determined by several values, namely 0.1 to 0.7 and the results are shown in Table 1.

Table 1 shows that the 0.1 and 0.2 threshold values of the segmentation image results are similar but some images do not show this so that the best threshold value is 0.1, while at the threshold value of 0.3 to 0.7 the detected objects are less and less perfect (intact). This is because some banana images have dark areas so that during the segmentation process the area value is smaller than the threshold value (T) so that during the segmentation process it will become a black image (0).

The next process is the feature extraction process which aims to take the characteristic value of the object so that this value becomes the determining parameter for classifying banana varieties. There are 7 features consisting of three-color features (Red, Green and Blue)

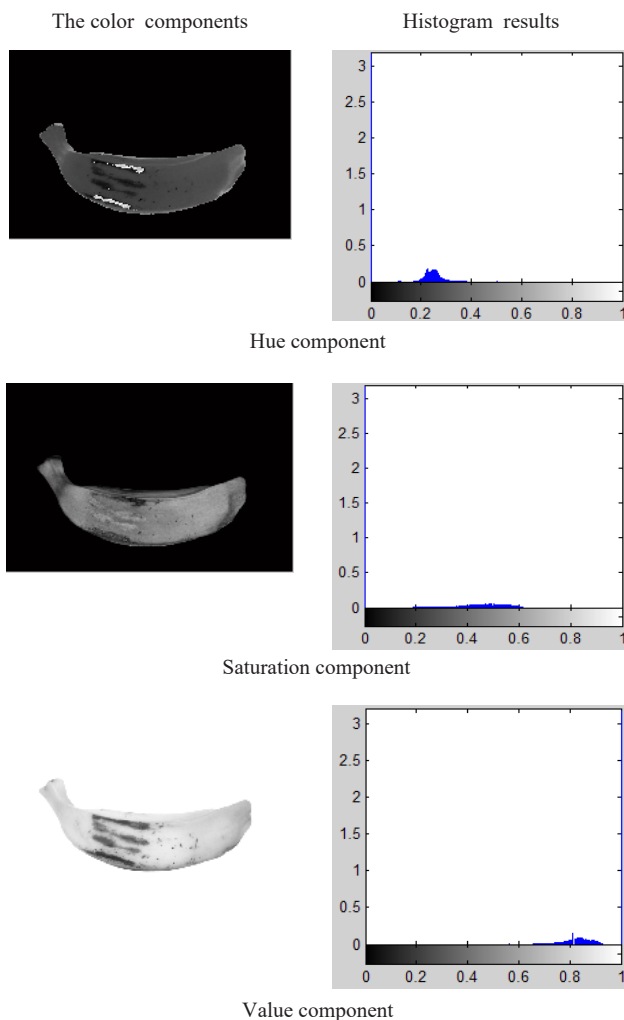


Figure 4. The difference in the image results of the HSV component is based on the results of the image histogram

Table 1. The results of the image segmentation based on the variation in the threshold value (T)

No	Threshold Value(T)	Image Segmentation Results
1.	0.1	
2.	0.2	
3.	0.3	
4.	0.4	
5.	0.5	
6.	0.6	
7.	0.7	

and four morphological features (area, perimeter, length and diameter). An example of the value of feature extraction is shown in Table 2.

Based on Table 2, it can be concluded that the perimeter parameter obtained values that are close to each other between the classes of raja banana, klutuk banana and merah banana. These parameters are part of the seven parameters used in the classification, so that 91.42% of the parameters are not close to each other. Based on this analysis, to classify banana varieties, it is enough to use general classification methods such as backpropagation and Radial Basis Function Neural Network. It is also different if the parameters of each class are close together, so it requires an advanced classification method.

The classification method used is backpropagation and radial basis function. The process of selecting training data and testing data is carried out randomly with the following

Table 2. The results of the feature extraction of banana varieties

Feature	Mas Kirana	Raja	Klutuk	Merah	Cavendis
Red	251	245	232	235	251
Green	249	237	244	222	250
Blue	243	211	225	219	240
Area	4051	14820	11362	10832	7274
Perimeter	303	589	550	508	439
Length	116	235	218	199	173
Diameter	39	71	63	60	42

Table 3. The results of system accuracy in the training process based on variations in the learning rate (α)

Variation of learning rate	Accuracy (%)
0.1	99.76
0.3	100
0.5	100

Table 4. Results of system testing accuracy based on variations in learning rate (α)

Learning rate	Accuracy (%)
0.1	96
0.3	88
0.5	98

conditions: 420 data is used for the training process and 50 data is used for the testing process. In the training process, the maximum number of iterations (epoch) is used of 500 with an error of 0.00000001. There is no limit regarding the amount of image data and the maximum iteration used in the training process because the principle of computer vision is that the more image data and iterations used in the training process, the system will be better at recognizing the pattern of the image. So, it is expected that the system will get better at distinguishing images in each class. The following are the results of the level of accuracy and the number of system training iterations based on the variation of the learning rate (α) as shown in Table 3. The number of test data is 50 and the accuracy results of system testing are shown in Table 4.

Table 4 shows that the test results from the test data get the highest accuracy of 98%. The data consisted of 50 banana data divided into 5 classes, each class containing 10 banana data. The lowest accuracy obtained from the testing process is 88%.

In this study, a comparison was made to other neural network methods such as the Radial Basis Function Neural Network. The results of the system testing accuracy resulting from the RBFNN method are 98.81% based on the confusion matrix calculation shown in Table 5.

Table 5 shows that in the Pisang Raja class, the total number of training data is 100 data, but during the training process the class was recognized as a pisang raja class with as many as 95 data and incorrectly recognized as a pisang Merah class with 3 data and 2 data as a pisang Klutuk class. The accuracy of the system in the process of testing of 100%.

The difference between the accuracy of the system in the process of training and testing can occur, because there is a difference of data on the training and testing process. Based on Table 2, several parameters in the pisang raja, pisang Merah and pisang Klutuk classes have close ranges of values, so it is possible for the system to misclassify the pisang raja class into another class. Table 6 shows a comparison of the system accuracy between the backpropagation method and the radial basis function method.

Table 5. Confusion matrix for the RBFNN training method

Output class					Target class
a	b	c	d	e	
100	0	0	0	0	a = Klutuk
0	100	0	0	0	b = Mas
0	0	100	0	0	c = Merah
2	0	3	95	0	d = Raja
0	0	0	0	20	e = Cavendis

Table 6. A comparison of the system accuracy between the backpropagation method and the radial basis function method.

Method	Training Accuracy (%)	Testing Accuracy (%)
Backpropagation	100	98
Radial Basis Function	98.81	100

IV. CONCLUSION

Based on the above research results, it can be concluded that the digital image processing process plays an important role in the classification process of banana varieties. There are two very important stages in digital image processing. first, the image segmentation process which aims to separate the research object from the background so that it will facilitate the feature extraction process. second, the feature extraction process which aims to extract the unique characteristics of each banana variety for the classification process.

To be able to be applied in real terms in the community, it is necessary to carry out further research by increasing the number of other banana varieties.

REFERENCE

- [1] Suyanti dan A. Supriyadi, *Pisang, Budi Daya, Pengolahan, dan Prospek Pasar*, XIX. Jakarta: Penebar Swadaya, 2008.
- [2] Kementerian Pertanian, *Outlook Komoditas Pertanian Sub Sektor Hortikultura Pisang*. Jakarta: Pusat Data dan Sistem Informasi Pertanian Kementerian Pertanian, 2016.
- [3] Badan Pusat Statistik dan D. J. Hortikultura, "Produksi Pisang Menurut Provinsi, Tahun 2015-2019," Jakarta, 2019.
- [4] C. P. Iklima dan M. Nasir, "Klasifikasi jenis pisang menggunakan metode k- nearest neighbor (KNN)," *J. Teknologi Rekayasa Informasi dan Komputer.*, vol. 1, no. 1, hal. 11-14, 2017.
- [5] Z. D. Lestari, "Sistem Klasifikasi Jenis Pisang Berdasarkan Ciri Warna HSV Menggunakan Metode K-NN," *Seminar Nasional Teknologi Informasi dan Komunikasi.*, hal. 11-15, 2019.
- [6] R. A. Pangestu, "Identifikasi Buah Pisang Berdasarkan Karakteristik Morfologi Buah dengan Metode Jaringan Syaraf Tiruan," thesis, Universitas Lampung, 2019.
- [7] L. R. Krisnadi, "Identifikasi Pisang Lilin, Mas Kirana dan Raja Sereh Menggunakan Jaringan Syaraf Tiruan," thesis, Universitas Panca Marga Probolinggo, 2019.
- [8] N. Rinanto, M. T. Wahyudi, dan A. Khumaidi, "Radial basis function neural network sebagai pengklasifikasi citra cacat pengelasan," *Rekayasa*, vol. 11, no. 2, hal. 118, 2018, doi:

- 10.21107/rekayasa.v11i2.4418.
- [9] Agusriandi, Elihami, dan W. Widiawati, "Identifikasi bawang merah dan bombay dengan pendekatan radial basis function neural network (RBFNN)," *J. Media Informatika Budidarma*, vol. 4, hal. 1043–1050, 2020, doi: 10.30865/mib.v4i4.2334.
- [10] Z. E. Fitri, I. K. E. Purnama, E. Pramunanto, dan M. H. Purnomo, "A comparison of platelets classification from digitalization microscopic peripheral blood smear," *2017 Int. Semin. Intell. Technol. Its Appl. Strength. Link Between Univ. Res. Ind. to Support ASEAN Energy Sect. ISITIA 2017 - Proceeding*, vol. 2017-Janua, hal. 356–361, 2017, doi: 10.1109/ISITIA.2017.8124109.
- [11] A. M. Nanda Imron dan Z. E. Fitri, "A classification of platelets in peripheral blood smear Image as an early detection of myeloproliferative syndrome using gray level co-occurrence matrix," *J. Phys. Conf. Ser.*, vol. 1201, no. 1, 2019, doi: 10.1088/1742-6596/1201/1/012049.
- [12] Z. E. Fitri, U. Nuhanatika, A. Madjid, dan A. M. N. Imron, "Penentuan tingkat kematangan cabe rawit (*capsicum frutescens* L.) berdasarkan gray level co-occurrence matrix," *J. Teknol. Inf. dan Terap.*, vol. 7, no. 1, hal. 1–5, 2020, doi: 10.25047/jtit.v7i1.121.
- [13] Z. E. Fitri, R. Rizkiyah, A. Madjid, dan A. M. N. Imron, "Penerapan neural network untuk klasifikasi kerusakan mutu tomat," *J. Rekayasa Elektr.*, vol. 16, no. 1, hal. 44–49, 2020, doi: 10.17529/jre.v16i1.15535.
- [14] Z. E. Fitri, A. Baskara, M. Silvia, A. Madjid, dan A. M. N. Imron, "Application of backpropagation method for quality sorting classification system on white dragon fruit (*Hylocereus undatus*)," *IOP Conf. Ser. Earth Environ. Sci.*, vol. 672, no. IT Agriculture, hal. 1–6, 2021, doi: 10.1088/1755-1315/672/1/012085.