Solanum Melogena Classification Using Features Shape and Color HSV Extraction with Backpropagation Neural Network Method

Muhammad Arbai, Abdul Syukur, Affandy, Moch Arief Soeleman and Wahyu Nofiyan Hadi

December 4, 2021

solanum melogena classification using features shape and color hsv extraction with backpropagation neural network method

[1]Muhammad arbai
Faculty of Computer Science Dian Nuswantoro University Semarang – Central Java email : muhammdarbai09@gmail.com

[2]Abdul Syukur
Faculty of Computer Science Dian Nuswantoro University Semarang – Central Java email : abah.syukur01@dsn.dinus.ac.id

[3]Affandy
Faculty of Computer Science, Dian Nuswantoro University, Semarang – Central Java email : affandy@dsn.dinus.ac.id

[4]Moch Arief Soeleman
Faculty of Computer Science Dian Nuswantoro University Semarang – Central Java e-mail : arief22208@gmail.com

[5] Wahyu Nofiyan Hadi
Faculty of Computer Science, Dian Nuswantoro University Semarang – Central Java
e-mail : navoleo7@gmail.com

Abstract-Eggplant (Solanoum Melongena L). is one of the commodities that have good prospects for development, besides that, eggplant is also rich in health benefits. The focus of this research is on the types of purple, green, round eggplant, and eggplant pokak because these four types have various benefits for body health and in general people think that every eggplant is the same, so classification is needed. to distinguish the types of eggplant and the nutrients contained in these vegetables. Besides that, many methods have been proposed to classify types of fruits and vegetables using digital images, but the results obtained are still below 84% accuracy, because previous studies only used one of the features of the HSV shape and color. In this research.

Keywords : Solanoum Melongena, Shape and Color Features of HSV, and BPNN

1. INTRODUCTION

Olericultural farmers who grow eggplant are generally sold based on a rough estimate in determining the type of eggplant. The determination of the price is generally through bargaining between the buyer and seller, not based on the type of eggplant sold. The problem that usually occurs is when there is no tool to determine the type and nutrition of eggplant, making it difficult for buyers or sellers to give prices *Solanoum Melongena*.

Research conducted by [2] using Hue Saturation Value (HSV) feature extraction and Local Binary Pattern (LBP) feature extraction got an accuracy of 65% for HSV feature extraction and 605 for LBP feature extraction. Another study conducted by [3] in which the image is placed under a five Watt lamp with a white cloth satirical system is able to acquire image data very well and for the hue color value all images are successfully obtained. For the results of testing the Naïve Bayes algorithm itself, the accuracy is 72.727%. According to research [4] using the technique of extracting shapes, colors and textures and the Learning Vector Quantization (LVQ) algorithm was successful with a score of 82%.

From several studies that have been mentioned using several feature extraction techniques and classification algorithms, in this study, feature extraction will be used The shape and color of HSV uses the Backpropagation Neural Network classification algorithm which is a variety of types of classification methods. This is done objectively by considering several relevant new studies.

2. THEORETICAL BASIS

A. HSV Color Features

Each color is used as a parameter obtained from several RGB (red green blue) color components.[5] The three components that color has as its main characteristics are as follows [6]

a. *Hue* serves to determine the color of red, green and so on.
b. *Saturation* function to determine color depth
c. *Value* is the brightness level of the color image

RGB to HSV color switch can use formula 1, formula 2, formula 4, formula 5 and formula 6 [7]

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

$$V = \max(r, g, b) \quad (2)$$

$$S = \{\frac{\max(r,g,b) - \min(r,g,b)}{\max(r,g,b)}, V > 0 \quad (3)$$

$$H = 60 * (\frac{(g-b)}{\max(r,g,b) - \min(r,g,b)}, jika\ V = r \quad (4)$$

$$H = 120 + 60 * (\frac{(b-r)}{\max(r,g,b) - \min(r,g,b)}, jika\ V = g \quad (5)$$

$$H = 240 + 60 * (\frac{(r-g)}{\max(r,g,b) - \min(r,g,b)}, jika\ V = b \quad (6)$$

$$H = \mathrm{H} + 360, \mathrm{jika\ H} < 0 \quad (7)$$

As for the HSV extraction features as follows [7]

a. *mean*

To determine the mean value can be seen in the formula

$$\mu = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N} I_{ij} \quad (8)$$

b. *Standard Deviation*

To determine the value of the Standard Deviation can be seen in the formula 6

$$\delta = \sqrt[2]{\frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N} \quad (I_{ij} - Mean)^2} \quad (9)$$

c. *Skewness*

To determine the value of the Standard Deviation can be seen in the formula 6

$$Skewness = \sqrt[3]{\frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N} \quad (I_{ij} - Mean)^3} \quad (10)$$

B. Shape Features

Some forms that can be used are the area of the object A, the perimeter of P, the compactness factor C and the roundness factor R. In using the shape image, the image is first converted into a binary image [4]

a. Perimeter

Perimeter is a calculation of the length of the boundary of the connected area. The adjacent pixels have two types, one of which is the adjacent pixels are in one parallel, namely the adjacent pixels are in the same column or the same row [5].

$$L = \sum_{i=0}^{n} \sqrt{(x_{(i+1)} - x_i)^2 + (y_{(i+1)} - y_i)^2} \quad (11)$$

b. Large

Area is the degree of similarity between the connected area and the minimum external rectangle. The computational area equation [6] is as follows;

$$R = \frac{S}{S_R} \quad (12)$$

c. Wide

Width is to describe the size of the connected region. Through image segmentation to get a binary value. Width is calculated using the following formula [7].

$$S = \sum_{(x,y)\in\Omega_i} f\,(x, y) \quad (13)$$

d. Diameter Ratio

Diameter is to get a measure of the ratio between the circumference denoted (P) and the diameter denoted by (D) is used for one of the features of the shape of the eggplant vegetable.

$$RD = \frac{P}{D} \quad (14)$$

e. Round

Roundness is to describe the size of non-separated pixels. Through binary image to get roundness value.

$$Roundness = \frac{4\pi A}{P^2} \quad (15)$$

f. Slim

Slenderness is a ratio comparison between the length of waste and vegetable bees[8]

$$Ramping = \frac{L_p}{W_p} \quad (16)$$

C. Neural Network

NN is a central processing unit designed to be distributed in parallel[8]

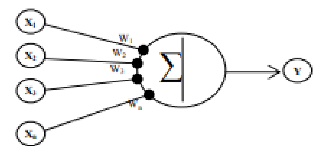

Figure 1 Communication between neurons

D. **Backpropagation Neural Network**

The requirement for the activation function in BPNN is that it must have a continuous nature, have derivatives or differentiate very easily and not a descending function. Purelin, tongsis and logsis are mandatory requirements to fulfill the activation function [9].

The following is the BPNN algorithm [10]

Process 0	:	Weight initialization Set maximum Epoh, Error Target, and Learning Rate Initialization Epoh = 0, MSE = 1.
Process 1	:	Perform the following process as long as Epoh < Maximum Epoh and MSE > Epoh Target Error = Epoh + 1
Process 2	:	Each Pair of elements to be trained, Do the next step

First phase: forward phase (Feed Forward)

Process 3	:	Each input unit () receives a signal and the signal is forwarded to all layers above it.$x_{i,}$ $i = 1,2,\ldots,n x_{i,}$
Process 4	:	Each unit in a hidden layer () adds up each weighted signal: Z_i $=x_{j,}$ $j = 1,2,\ldots,p n_i b1_j + \sum_{1=1}^{n} X_i V_{ij}$ Use the activation function to calculate the output signal = f(Z_) and send the signal to all units in the upper layer $Z_j in_j$
Process 5	:	Each output unit adds up the weighted input signals:

$$Y\_in_k = b2_k + \sum_{i=1}^{p} Z_i W_{jk}$$

Use the activation function to calculate the output signal and send the signal to all units in the upper layer $Y_k = f\ (Y\_in_k)$

The second phase: the backward phase (Back Propagation)

Process 6 : Each input output unit is the sum of the units above it;$Y_k\ (j = 1,2, \dots, p\ )$

$=()\ ()t_k - Y_k f' Y\_in_k$

$$\varphi 2_{jk} = \varphi_k Z_j$$

$$\beta 2_k = 5_k$$

The next step is to calculate the weight correction

$=W_{jk} \alpha \varphi 2_{jk}$

After completing the weight correction calculation, the next process is to calculate the bias value

$=b2_k \alpha \beta 2_k$

Process 7 : Each hidden unit adds up its input delta from the different units in the layer above it; $Z_j\ (j = 1,2, \dots, p\ )$

$$\delta in_j = \sum_{k=1}^{m} \delta 2_k W_{kj}$$

Multiply this value by the derivative of the activation function to calculate the error information:

$\delta 1_j = \delta\_in_j f'(Z\_in_j)$

$\delta 1_{ij} = \delta 1_j X_j$

$\beta 1_j = \delta 1_j$

The next process calculates the weight value

$\triangle V_{ij} = \alpha \varphi 1_{ij}$

The next step is to calculate

$\triangle b_{ij} = \alpha \beta 1_{ij}$

Process 8 : Each output unit corrects the bias and its weight (j = 0.1,2,..,p)$Y_k\ (k = 1,2, \dots, m\ )$

$W_{jk}(baru) = W_{jk}(Lama) + \triangle W_j$

$b1_j(baru) = b1_j(Lama) + \triangle b_j$

Each hidden unit fixes the bias and its weight (i = 0.1,2,…,n)$Z_j\ (j = 1,2, \dots, p\ )$

$V_{ij}(baru) = j_{ij}(Lama) + \triangle V_{ij}$

$b1_j(baru) = b1_j(Lama) + \triangle b_j$

E. Framework

The background of this research is to classify the types of eggplant in Nganjuk, East Java, this research is the first on the classification of eggplant species in digital image processing. Here's a schematic framework.

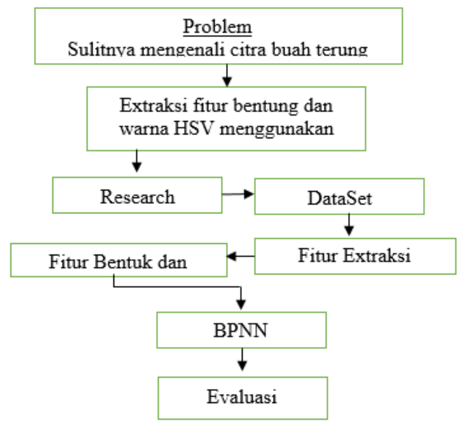

Figure 2 Research Framework

3. RESEARCH METHODS

In this study, digital images of eggplant were used by using 4 types of eggplant varieties. The dataset is as follows.

Picture 3 Pokak Eggplant

Picture 4 Purple Eggplant

Figure 5 Round Eggplant

Figure 6 Green Eggplant

Here is the proposed method

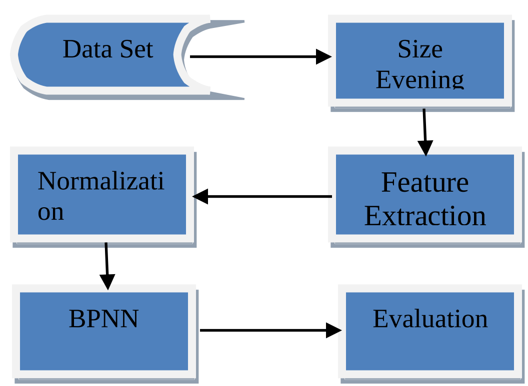

Figure 7 Proposed Method

The feature extraction used is the shape features including perimeter, area, width, diameter, round, slender, while the color features taken include mean, skewness and standard deviation.

4. RESULTS AND DISCUSSION

A. Data Collecting

The total data set used was 400 eggplant images consisting of four types, pokak eggplant = 100, purple eggplant = 100, round eggplant = 100, and green eggplant = 100. The digital image processing step is carried out in order to produce a value so that it can be processed with data mining .The first step is to separate the object from the background, like the image below

Figure 8 Original Image

Figure 9 Preprocessing Results

The next stage after preprocessing the image then the image resulting from the background separation is processed and converted into a binary image, but before the binary image processing step it is necessary to process the gray image change. And the last process is to find the value of the HSV shape and color features.

Figure 10 Image of gray

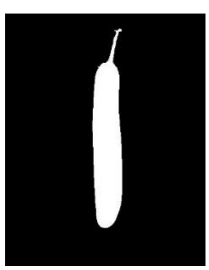

Figure 11 Binary Image

B. Test

Attribute 9 is used which is divided into two features, namely HSV shape and color features, while the attribute features are mean, skewness, standard, perimeter, area, diameter, round, slim and wide. After all attribute values are obtained, the testing process is carried out using a rapid miner. The tests carried out determine the learning rate value with a range of 0.1 - 0.10 momentum 0.1 - 0.10, validation number 1 - 10 and training cycles 200. The most optimal parameter value learning rate 0.1 momentum 0.9 training cycles 200 validation number 7 results obtained accuracy = 99.75%, recall = 99.75%+/-0.75% and class precision = 99.77%+/-0.68%.

	true TB	true TP	true TPU	true TU	class pprecision
pred. TB	100	0	0	0	100.00%
pred. TP	0	100	0	1	99.01%
pred. TPU	0	0	100	0	100.00%
pred. TU	0	0	0	99	100.00%
class recall	100.00%	100.00%	100.00%	99.00%	
Mean Recall	99.75%+/-0.75%				
Mean class precision	99.77%+/-0.68%				

Figure 12 Calculation Results

5. CLOSING

From the results of experiments that have been carried out, the use of HSV shape and color feature extraction and the BPNN method obtained results with a learning rate of 0.1 momentum 0.9 training cycles 200 validation number 7 results obtained accuracy = 99.75%, recall = 99.75% +/-0.75% and class precision = 99.77%+/-0.68%. Suggestions for further research is to increase the number of data sets and or the number of attributes to make it more varied in the classification process.

BIBLIOGRAPHY

[1] BPS., 2017. The role of the agricultural sector in regional development. Bogor.

[2] Ayuningsih rahmat et al, Classification of Food Image Using HSV Color Moment and Local Binary Pattern with Naïve Bayes Classifier vol 3, 2019.

[3] Andy Susilo Faizal et al, “Classification System of Frozen Tuna Fish Quality Based on HSV Color Value Features Using Naïve Bayes Method,” , vol. 3, hh. 753–760, 2019.

[4] Sutarno et al, "Identification of Vegetable Plants Based on Features of Shape, Color and Texture of Leaves Based on Image Processing and LVQ," vol. 3, 2017.

[5] Muljono, “Digital Image Processing,” Andi Publisher, Yogjakarta, 2017

[6] Qingfeng Zhou et al, *Analysis of Shape Features of Flame and Interference Image in Video Fire Detection,* IEEE 978-1-4673-7189-6/15, 2015.

[7] J. Zhou, X. Du,"*Application of image recognition technology in fire detection system of ship*," Fire Science and Technology, vol. 26, no. 4. pp.417-420, 2007.

[8] L. Bo, X. Yi, Z. Lin, “Image Characteristic Analysis of Moving Fire Diffusion Flame in Circular Motion,” Journal of Combustion Science and Technology, vol. 19, pp. 60-67, 2013.

[9] Wu, Qingfeng, Changle Zhou, and Chaonan Wang, (2006). “Feature extraction and automatic recognition of plant leaf using artificial neural network”. Advances in Artificial Intelligence 3.

[10] Junhua, C. & Jing, L., 2012. *Research on Color Image Classification Based on Hsv Color Space. Second International & Measurement Computer, Communication and Control*

[11] The gift of Ayu Ningsih, et al., *"Classification of Meaningful Images Using HSV Color Moment and Local Binary Patterns with Naïve Bayes Classifier"*, e-ISSN:2548-964X,Vol 3, 2019, p. 3166-3173

[12] Singh, SM & K., H., *"Content-Based Image Retrival Using Color Moment and Gabor Texture Feature"*. IJCSI International Journal of Computer Science Issues, 2012, pp. 399-408.

[13] *Rajasakeran* S., GA. Vijayalakshmi Pai, "Neural Network, Fuzzy Logic and Genetic Algorithms", prentice-Hall of india, New

[14] Heaton, J,"*Introduction to Neural Networks With Java*" Publisher: Heaton Research. ISBN: 1-60439-008-5