The Sorting and Grading of Red Chilli Peppers (*Capsicum annuum* L.) Using Digital Image Processing

Nafis Khuriyati[1], Agung Putra Pamungkas[2], Anggraito Agung P.[3]

Department of Agroindustrial Technology, Universitas Gadjah Mada
Flora Street No.1 Bulaksumur, Yogyakarta, Indonesia

Abstract

Sorting of red chilli pepper manually makes the result unobjective and inconsistent. By using digital image processing and Artificial Neural Network (ANN), we can minimize the subjectivity and inconsistency. The aim of this research was to design ANN architecture for sorting and grading red chilli pepper based on image analyses. Red chilli pepper cv. Helix image was processed using digital image processing and classified using ANN. Images from the two sides of a red chilli that were 20 samples taken in a closed box equipped with a smartphone. The images of red chillies were processed by using MATLAB application that was used to transform RGB image to obtain mark of colour features and texture features. Subsequently, colour and texture feature were chosen to get variables for composing ANN. It was discovered that red, energy, and correlation are variables to compose ANN. This ANN had 3 input cells, 22 hidden layers cells, and 5 output cells and could classify 20 samples of red chilli peppers with 84.46% level of accuracy.

Keywords — *Artificial neural network, digital image processing, grading, red chilli peppers, sorting*

I. INTRODUCTION

The chilli pepper (*Capsicum* annuum) is the most widely grown and consumed spicy crop around the world, particularly in Asian and South American countries. It is estimated that the global production is approximately 40.7 million tons, and the area harvested was approximately 3.8 million ha in 2017 [1]. The chilli pepper is one of important commodity for Indonesia's economy. From consumption side, chilli has big share reflected from inflation weight reach 0.35%. The chilli is also used as flavoring for most processed products like instant noodle and chilli sauce with high product value [2].

Sortation is part from post-harvest that do with aim to separate post-harvest result that good and bad. While grading do with classification chilli based on some physical properties. That is done to conform to the market demand. Consumers of household, food industry, hotel, supermarket, and restaurant have different specifications. Sortation and grading are done by traders, while farmers generally sell chillies directly to traders or collectors [3]. Indonesian national standard for fresh red chilli right now referring to SNI (Standar Nasional Indonesia) 4480:2016. The red chilli is considered fresh chillies when they are ripe and reddish, in fresh condition, not rotten, not deformed or damaged, and free from pest and disease. Based on the percentage amount of defect, chillies are classified into three quality levels i.e., super, the first, and second class.

Today, sortation and grading of red chillies are still done in manually through visual observation. This method has some weaknesses, such as that it produces inconsistency due to the limited capacity of human visual sense, fatigue, and the divergence of quality perceptions from different observers [4]. Determining the chillies' quality after sorting all of them takes a long time, and this can also affect the quality of the chillies after harvesting. This must be anticipated by applying a method and using tools for initial decision makers before the chillies are properly sorted. For this reason, a faster post-harvest handling technology with a low error rate is needed to help equate or standardize perceptions of the quality of sorting and chili grading among the parties involved. One alternative technology that can be used is image processing; this image method is a non-destructive (without damaging materials) quantitative method [5]. Digital image processing can be used for factors, such as shapes, sizes, and color recognition, texture extraction, body strength, smell and handling [6]. Mechanization, rapid screening, and reduction of human error are considered as the main advantages of the image processing [7], [8] which has been emphasized in industries and factories [9]. With the development of this alternative, it is hoped that it can be a system of early decision making in determining the quality of red chillies. In addition, the buyer also would not need to wait for the sorting to be carried out by the seller to find out the quality of the chillies to be sold.

The technology to detect the level of quality and sorting has been developed to help obtain maximum results and with a short period of time. This technology has been developed by developed countries to be applied in sorting machines and quality control of agricultural products. However, in Indonesia this technology has not been widely used for the purposes of sorting and grading horticultural products because the equipment is still relatively

expensive. Other technologies that can be applied to automatic sorting and decoding machines are digital image processing (Image Processing). Good image processing can be obtained when combined with a decision-making system that is able to provide results with high accuracy. This technology is cheaper because it only requires a charge couple device (CCD) camera and a computer equipped with frame grabber images, which are already widely available in the market [10].

In recent years, there has been a growing interest among researchers to develop image processing applications using a smartphone camera [11], [12]. This could be attributed to the fact that the prices of smartphones are generally low while their processing capabilities are enhanced thanks to advanced processors, high-resolutions cameras, and memory storage devices [13].

In a pattern recognition system, ANN can identify the distinguishing features needed to carry out pattern recognition tasks. After the ANN is given a training set in the form of a number of patterns, the network will classify objects according to the patterns in certain categories [14]. This can be applied also in sorting red chillies. Because every red chili harvested and sold has a pattern or characteristic that can be read by ANN, the aims of this research were to identify relations between quality of red chilli peppers with colours and texture elements using digital image processing method and to design an ANN architecture to perform red chilli pepper sortation based on image analyses.

II. MATERIALS AND METHOD

A. Size of Dataset

The study was carried out on red chili peppers cv. Helix which were picked from the plantations of farmers in Kulon Progo Regency, Indonesia. A total sample of 478 chilli peppers consisted of normal, green (immature), broken, rotten, and moldy chillies. This sample was then grouped into five categories in accordance with SNI. The sample was divided into two data sets that were training and validation sample set. The training set consisted of 190 chillies (38 chillies from each category), and the validation set consisted of 288 chillies (235 normal, 21 green, 9 broken, 7 rotten, 16 moldy chillies). The data needed in this study were the colors (red, green, blue, hue, saturation, value) and textures (entropy, energy, correlation, contrast, homogeneity) of the sample. The tools used were a photo box measuring 60 x 40 x 48 cm, 2 pieces of USB LED light strip, a power bank, a smartphone, a laptop, and MATLAB 2014b application (Fig 1.)

B. Capturing and Processing Chilli Images

Image capturing was conducted using a Xiaomi Redmi Note 4 smartphone with a 13 MP f / 2.0 camera with manual mode settings and with ISO 100 and white balance auto mode in a closed box with a black background. There were 20 chillies as shown in Fig 2. The image was taken twice and then moved to the laptop for image processing using MATLAB application. Subsequently, one set of image was cropped using MATLAB application manually as shown in Fig 3. For the results of the image of chili being cropped using MATLAB automatically and manually can be seen in Fig 4. After all sets of images were cropped, these images were processed using MATLAB application. To simplify the processing, a Graphic User Interface (GUI) was made, which could display the original image and the results of image processing both in the form of images and numbers that can be seen in Fig 5. The stages in the image processing GUI were compiled in the form of taking sample images, converting RGB images to HSV images, converting HSV images to binary images, thresholding and image segmentation, returning images to RGB colors, extracting color features and textures.

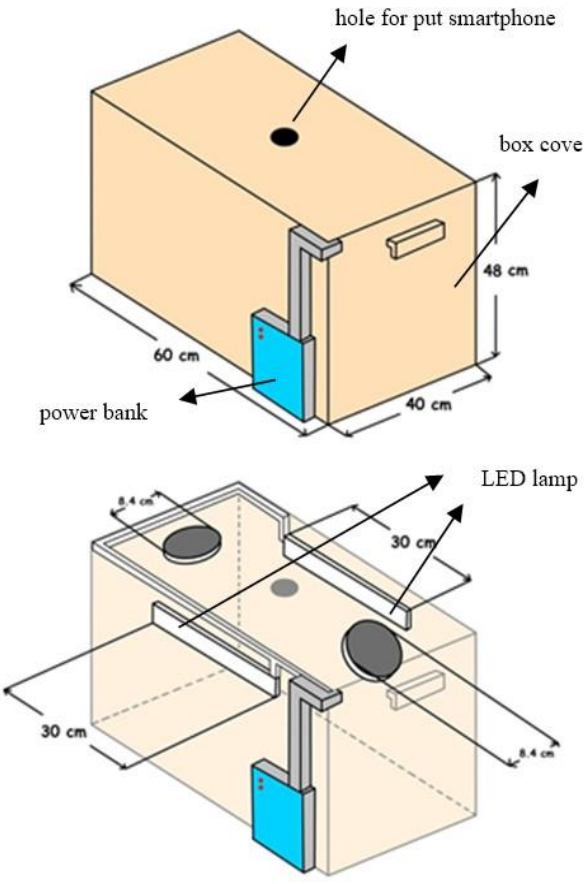

Fig. 1: Tool for taking some images of chillies

Fig. 2: One set of chili images

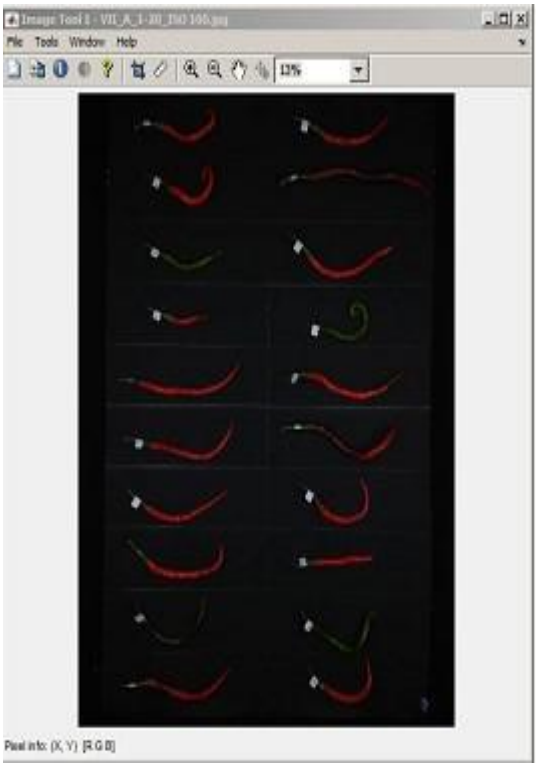

Fig. 3: Application for crop images

(a) (b)

Fig 4: Results of cropping of chilli images through manual method (a) and automatic method (b)

III. RESULTS AND DISCUSSION

A. System Performance

After all image data were stored in an excel file, the data uniformity was tested. This was conducted to find out whether all image data were uniform or not. Uniform data are within the Lower Control Limit (LCL) and within the Upper Control Limit (UCL). Out-of-control data were not used in determining ANN parameters.

After removing out-of-control data, uniform image data were used to determine the parameters that were used to design the ANN. To determine the parameters of the ANN, correlation analysis was carried out. The purpose of the use of correlation analysis was to determine the level of closeness of the relationship between variables expressed by the correlation coefficient. Correlation analysis was carried out by calculating the correlation value of the characteristic variable of red chili consisting of 5 characteristics of chili (normal, green, broken, rotten and moldy) and 11 parameters resulting from image analysis, namely red, green, blue, hue, saturation, value, entropy, contrast, homogeneity, energy and correlation. Calculation of correlation analysis was performed by using the SPSS program.

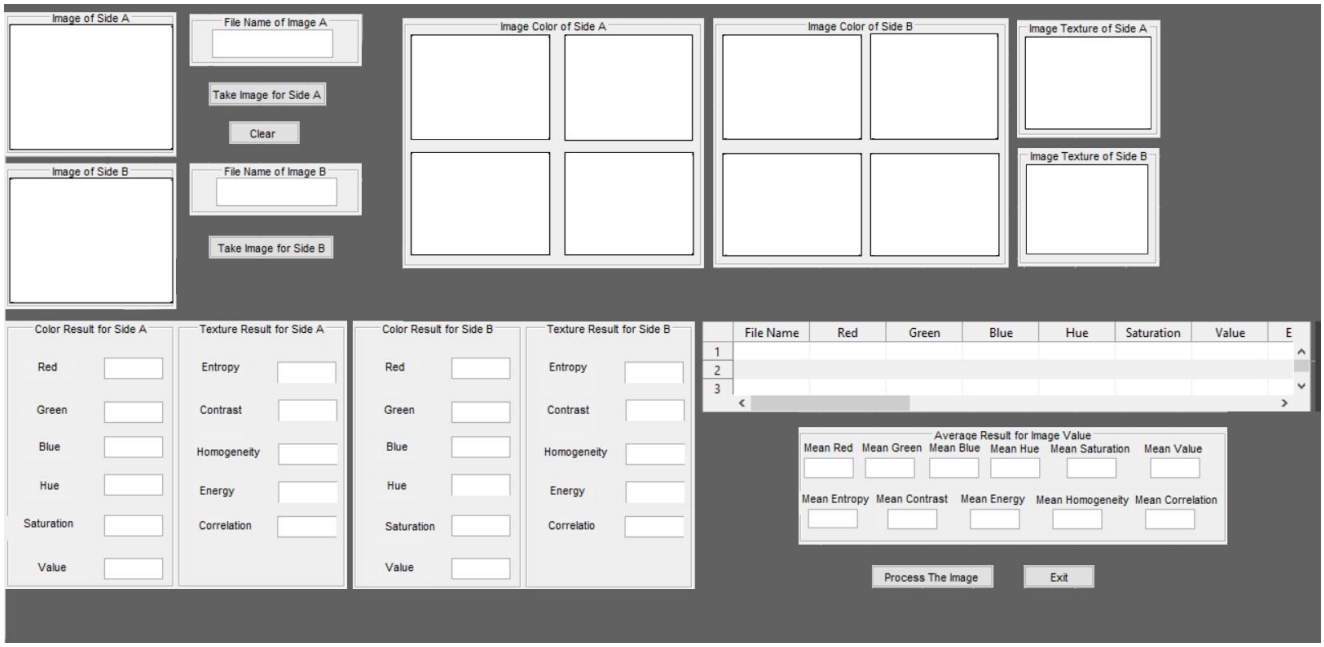

Fig. 5: GUI for image processing

The parameter chosen is a parameter that has a correlation value of ⩾0.2. The determining of this correlation value was based on research conducted by [15], where the selection of variables with correlation values was able to correctly classify the maturity level of bananas. Based on the results of the correlation analysis carried out on the variables of the characteristics of processed chili images, the colour red, energy, and correlation variables met the criteria for the correlation value determined. This showed that the three variables had a degree of closeness among the five characteristics of the chili studied. The colour features that represent the influential variables on the characteristics are red. Red variable here had an influence because it could distinguish the characteristics of the chillies which had a color difference, such as between normal chillies and green chillies. While the texture features that represent influential variables on characteristics are energy and correlation. Energy features were used to measure uniformity or commonly called angular second moments. Energy will be of high value if the pixel values in the image are similar to each other. Conversely, if the energy value is small, the pixel values are heterogeneous/not the same image. The energy feature is useful for distinguishing the characteristics between normal chillies and broken as well as between rotten and moldy chillies. For other texture features, such as correlation, they have the highest correlation value between red and energy. This indicates that this correlation feature has a considerable influence on distinguishing the characteristics of chili because correlation was used to measure the linearity of a number of pairs of pixels in the image. Table 1 below shows the results of the correlation analysis.

Table I. Results of Correlation Analysis of Color and Texture Features

Parameter	Correlation Value	Significance Value
Red	**0.279**	**0.000**
Green	0.143	0.005
Blue	0.047	0.358
Hue	0.101	0.046
Saturation	0.046	0.360
Value	0.143	0.004
Entropy	0.152	0.003
Contras	0.081	0.109
Homogeneity	0.090	0.074
Energy	**0.226**	**0.000**
Correlation	**0.329**	**0.000**

B. Compilation and Training of ANN

The designing of ANN started with the selection of training methods, the value of learning rate, the momentum value of constants, the number of hidden layer cells, and the activation function. Measurements in getting the performance of the design can be seen from the value of mean square error (MSE). The selection of ANN model can be seen from the MSE value, the number of iterations, and the

accuracy of identification. Preparation of ANN was done using MATLAB software and the neural network toolbox. The design of ANN for sorting consisted of 3 neural layers, namely input layer, hidden layer, and output layer. The input layer consisted of 3 layers which were represented by predetermined parameters (red, energy and correlation).

Table II. Results of ANN Validation with 288 Samples

Validation Image	Sorting Results					Sample Total	Sample Accuration	Average Accuration
	Normal	Green	Broken	Rotten	Moldy			
Normal	227	1	4	2	1	235	96.59%	82.98%
Green	0	12	1	5	3	21	57.14%	
Broken	8	1	0	0	0	9	0%	
Rotten	7	0	0	1	1	7	0%	
Moldy	11	0	1	4	0	16	0%	

Table III. Results of ANN Validation with 55 Samples

Validation Image	Sorting Results					Sample Total	Sample Accuration	Average Accuration
	Normal	Green	Broken	Rotten	Moldy			
Normal	7	1	2	1	0	11	63.64%	21.81%
Green	0	4	1	4	2	11	36.36%	
Broken	10	0	0	1	0	11	0%	
Rotten	9	0	1	1	0	11	9.09%	
Moldy	6	1	2	2	0	11	0%	

Hidden layers were determined by trials and errors to get a design that had maximum performance. On the other hand, the output layer consisted of 5 layers in the form of 5 characteristics of chili studied (normal, green, broken, rotten and moldy).

After obtaining the ANN design training method with a low MSE value, the ANN design was obtained which was then used for sorting and grading red chillies. The following is an arrangement of ANN designs that was used:

1) Algorithm of training: trainrp
2) Learning rate value: 0.6
3) Momentum constant value: 0.6
4) Number of hidden layer cells: 22
5) Function of input activation: logsig
6) Function of hidden layer activation: logsig
7) Function of output activation: logsig
8) Maximum value of iteration (epoch): 10000
9) Value of Mean Square Error: 0.001

C. Validation of ANN and Making GUI

After network training was completed, validation of the network was done. The validation of the network was done to show the extent of the network's ability to read patterns similar to the training process. Network validation used weight values and training results bias, where both values stored the results of training conducted by the network. The validation was done using different chilli sample data. The validation data used were 288 chillies, which included red, energy, and correlation values. Based on the results of the validation in Table 2, of the 288 validated test data, there were 239 correctly identified data and 49 incorrectly identified data. The obtained results of accuracy was 82.98%. Chillies with broken, rotten, and moldy characteristics could not be recognized properly by the system. This was because the physical characteristics of broken chillies tended to be similar to normal chillies. For chillies with rotten and moldy characteristics, the system had not recognized them maximally because the color of the rotten and moldy chillies tended to resemble that of normal chillies. Apart from the similarity factor to one of the chili peppers, other factors that could cause the system not to work optimally were the lack of broken, rotten, and moldy chili samples during data collection in the field. Therefore, the system still had difficulty learning and reading the images of broken, rotten, and moldy chillies. Nevertheless, the result of the data validation that used 55 chillies obtained 21.81% accuracy with 12 correctly identified data and 43 incorrectly identified data. This significant difference was due to the accuracy of each characteristic of the chillies. In table 2, normal and green characteristics had an accuracy of 96% and 57%. In table 3, normal, green, and broken characteristics have some accuracies of 63%, 36%, and 9%. The accuracy of normal characteristics in table 2 affected the level of accuracy for the sample in table 2 as a whole. This was due to differences in the number of samples used for validation. Therefore, the quantity of samples that were trained and validated was directly proportional to the quality of the network itself. According to [16], small dark marks and defects to the edges of images are difficult to detect. Some improvements were made through changing the geometry of the tool, for example, increasing image resolution to capture more minor defects or allowing rotation in two fields, which could eliminate problems with edge defects. To sort the tested chillies, ANN that were built already had the ability to distinguish normal chillies from green chillies. The ability of the camera of a smartphone will be more optimal if it is used to sort chilli samples or

other agricultural product samples based on the color features of the sample images.

In sorting red chillies by using ANN, a user interface was needed to display the results of sorting and grading easily and clearly. To facilitate the sorting and grading of red chillies, the graphic user interface (GUI) feature was used. The GUI for sorting and grading red chillies consisted of several image columns to include the image of the chillies and

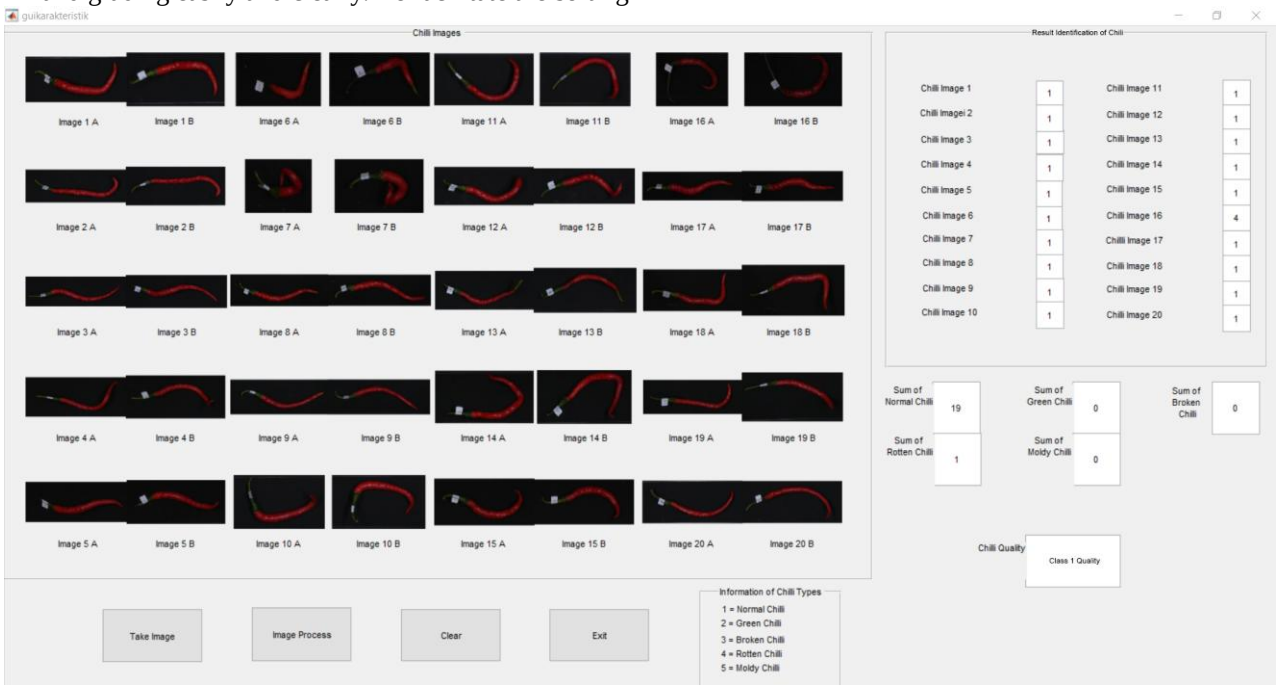

Fig. 6: Results of sorting and grading using a GUI

several columns in the form of information about the quality of each chilli that was analyzed.

To operate the GUI sorting and red chilli grading, first press the "take picture" button to load 40 images of chillies. Then, proceed with pressing the "image process" button to start the sorting process and red chilli pepper grading. Finally, the results of the simulation will be displayed in several columns in the GUI. The results of chili sorting are numbers which have their own information according to the state of each image that has been loaded. Examples of these can be seen in Fig 6 that shows the results of the quality of 20 chilies which are loaded with their images.

IV. CONCLUSIONS

The features of color and texture used for sorting red chillies were the colour red, energy, and correlation features. The colour red, energy, and correlation features had a correlation values of 0.279, 0.226, and 0.329. The artificial neural network architecture that was used had 3 layers in the form of an input layer with the colour red input, energy, and correlation, a hidden layer with 22 hidden cells, and an output layer with output in the form of normal chillies, green chillies, broken chillies, rotten chillies and moldy chillies. The results of the validation of ANN showed an accuracy rate of 84.46%.

ACKNOWLEDGMENT

This research was financially supported by Ministry of Research, Technology, and Higher Education of the Republic of Indonesia through a research grant for Penelitian Terapan Unggulan Perguruan Tinggi-Universitas Gadjah Mada 2018 (Leading Applied Research of Higher Education-Universitas Gadjah Mada).

REFERENCES

[1] FAO, 2017. Food and Agriculture Organization of the United Nations Statistic Division website. [Online]. Available: http://www.fao.org.

[2] Farid, M. & Subekti, N. A., 2012. Tinjauan terhadap Produksi, Konsumsi, Distribusi dan Dinamika Harga Cabe di Indonesia. Buletin Ilmiah Litbang Perdagangan, VOL. 6, No. 2: 211-234.

[3] Alif, S. M. 2017. Kiat Sukses Budidaya Cabai Keriting. Yogyakarta: Bio Genesis.

[4] Gunasekaran, S. 1996. Computer Vision Technology for Food Quality Assurance. Trends in Food Science & Technology, 7 (8), 245-256.

[5] Sandra. 2005. Aplikasi jaringan syaraf tiruan untuk pendugaan mutu mangga megar mecara non-destruktif. Jurnal Teknologi Pertanian,Vol. 6 No. 1 (April 2005) 66-72

[6] Meruliya, T., Dhameliya, P., Patel, J., Panchal, D., Kadam, P., Naik, S., 2015. Image Processing for Fruit Shape and Texture Feature Extraction-Review. International Journal Comput. Appl. (0975-8887) Vol. 129, No. 8: 30-33.

[7] León, K., Mery, D., Pedreschi, F., León, J., 2006. Color measurement in L∗a∗b∗ units from RGB digital images. Food Res. Int. 39, 1084–1091

[8] Payne, A.B., Walsh, K.B., Subedi, P.P., Jarvis, D., 2013. Estimation of mango crop yield using image analysis – segmentation method. Comput. Electron. Agric. 91, 57–64.

[9] Shamili, M., 2019. The estimation of mango fruit total soluble solids using image processing technique. Scientia Holticulture Vol. 294: 383-389.
[10] Ahmad, U. 2010. Aplikasi teknik pengolahan citra dalam analisis non destruktif produk pangan. Jurnal Pangan Vol. 19 No. 1 Maret 2010.
[11] Chung, S., Bereshears L., E., Yoon, J., 2018. Smartphone Near Infrared Monitoring of Plant Stress. Comput Ekectron Agriculture Vol. 154: 93-98
[12] Shrivastava, S., Lee, W., Lee, N., 2018. Culture-free, highly sensitive, quantitative detection of bacteria from minimally processed samples using fluorescence imaging by smartphone. Biosens. Bioelectron. Vol. 109: 90–97.
[13] Hosseinpour, S., Ali H. I., Mortaza., A., 2019. An Intelegent machine vision-based Smartphone App for Beef Quality. Journal Food Engineerig 248: 9-22.
[14] Hermawan, A. 2006. Jaringan Saraf Tiruan Teori dan Aplikasi. Yogyakarta: Penerbit Andi.
[15] Widjanarko, D. 2014. Program Aplikasi Untuk Identifikasi Kematangan Buah Pisang Mas (Musa Paradisiaca l) Berdasarkan Parameter Citra dengan Teknologi Pengolahan Citra Digital. Skripsi Teknologi Industri Pertanian. Universitas Gadjah Mada: Yogyakarta.
[16] Bennedsen, B. S., D. L. Peterson and A. Tabb. 2007. Identifying apple surface defects using principal components analysis and artificial neural networks. American Society of Agricultural and Biological Engineers Vol. 50(6): 2257-2265J.