Designing the quality of coffee bean detection application using Hue Saturation Intensity

T H Nasution\(^1\)*, M Rumansa\(^1\), Lukman Adlin Harahap\(^2\)

\(^1\)Department of Electrical Engineering, Universitas Sumatera Utara, Medan, Indonesia
\(^2\)Department of Mathematics and Computer Science, Faculty of Science, Chulalongkorn University, Bangkok, Thailand

*Email: tigor.nasution@usu.ac.id

Abstract. This study discusses the design of applications that can detect the quality of raw coffee beans both good and bad based on the value of HSI (Hue Saturation Intensity) in the coffee beans using digital image processing and Backpropagation artificial neural networks. The process of identifying coffee beans is based on the intensity of the HSI value that is owned by coffee beans. The HSI value is converted from RGB values, then the training process is performed on backpropagation artificial neural networks to recognize good quality seeds and poor-quality seeds. The testing phase is carried out using an interface designed in the MATLAB R2013a software. Based on the results of testing, it was found that this application was able to detect samples of coffee beans that were not properly trained.

1. Introduction

Coffee is one of the superior agricultural commodities\(^1\)–\(^3\). The economic value of coffee beans is strongly influenced by its quality. However, not all coffee beans produced by coffee farmers are of good quality. Coffee beans can be damaged due to several things such as being attacked by pests or because of the post-harvest process, for example the process of stripping fruit skin is less than perfect and the process of washing and drying is not evenly distributed. This causes some parts or all parts of the coffee beans to become damaged. Therefore, it is necessary to select quality coffee beans before making a sale to increase the selling value of the coffee beans.

Generally, the selection of quality coffee beans is still done manually by visually looking at coffee beans\(^4\). In this way, the selection of good coffee beans can only be done by people who are knowledgeable about the characteristics of good coffee beans. Based on these reasons, in this study we designed a software to recognize good quality or damaged coffee beans. With this software ordinary people will be able to distinguish the quality of coffee beans from the visual without needing to know about the characteristics of coffee beans that are visually good.

This research is based on several previous studies about the quality of coffee beans. Previous studies have been carried out to evaluate the quality of coffee beans using infrared spectroscopy\(^5\). In addition, there is also the introduction of good coffee beans using image processing based on RGB color components\(^6\). We have also conducted research using image processing to determine the type of coffee roasted\(^7\). In contrast to some previous studies we will design software to recognize the quality of coffee beans by using image processing and artificial neural networks based on the intensity of HSI (Hue Saturation Intensity) values.

Coffee beans can be damaged due to several things such as being attacked by pests or because of the post-harvest process, for example the process of stripping fruit skin is less than perfect and the process...
of washing and drying is not evenly distributed. This causes some parts or all parts of the coffee beans to become damaged. Damage to coffee beans can be detected, one of them by looking at the spread of colors that are less uniform in coffee beans. Visually poor-quality coffee beans generally have black patches and tend to have dull and dry skin on the surface. This is very different from good quality coffee beans, where the surface of the skin looks fresher and has a more evenly spread color on its surface as shown in Figure 1.

![Figure 1. Good quality and bad quality coffee beans](image1)

In this research, good and bad quality coffee beans will be detected based on the color distribution. Based on SNI 01-2907-2008[8], regarding the determination of the value of defective coffee beans, several parameters will be used which can determine a poor-quality coffee bean based on the distribution of colors such as: full and partial black seeds, full brown beans and some as shown in Figure 2.

![Figure 2. Black and brown coffee beans](image2)

2. Method

This research stage begins with the acquisition of the image of the beans, and then process the data pre-processing resulting features/characteristics that can distinguish the quality of the coffee beans. Then the training phase is done using backpropagation artificial neural networks. At the last stage, the creation of an interface makes it easier in the testing process. The stages of this research flow can be seen in Figure 3.
2.1. Image acquisition
In the image acquisition stage, the image of coffee beans is taken as training material. Image capture is done using the Xiaomi Redmi Note 5 smartphone camera with 12MP f1/9 resolution with lighting using LED lights. Examples of image acquisition results can be seen in Figure 4.

![Example image](image1.jpg)

Figure 4. Example image

2.2. Pre-processing
Pre-processing starts with cropping and normalizing the size to uniform image size. This is also done because the resolution of the image produced by the camera is large enough to slow down the subsequent processing. In the cropping process the image size is uniformed to 150 x 150 pixels as shown in Figure 5.
The next step is to get the HSI value from the image RGB value. Previously the RGB value of the image was normalized first from 0-255 to 0-1. This is because the range of saturation and intensity values in the HSI ranges from 0-1. The following is the equation used to convert the values of R, G, B to H, S, I.

\[
\begin{align*}
\theta &= \cos^{-1}\left(\frac{\frac{1}{2}(R-G)+(R-B)}{\sqrt{(R-G)^2+(R-B)(G-B)^2}}\right) \quad (1) \\
H(\text{Hue}) &= \begin{cases} 
\theta & \text{if } B \leq G \\
360 - \theta & \text{if } B > G 
\end{cases} \quad (2) \\
S(\text{Saturation}) &= 1 - \frac{3}{R+G+B} \min(R,G,B) \quad (3) \\
I(\text{Intensity}) &= \frac{1}{3}(R + G + B) \quad (4)
\end{align*}
\]

The Hue value that ranges from 00 - 3600 is also normalized so that the value becomes 0-1 with the equation \(H = \frac{H}{360}\). The saturation value that has been obtained is used to separate objects (foreground) in the background using the thresholding technique.

The next process is the thresholding process. Thresholding is used to separate objects (foreground) from the background[13,14]. This process is done by setting a threshold value, where the RGB value that is equal to or more than the threshold value will be changed to white (1), while the RGB value that is less than the threshold value will be changed to black (0). In this case the threshold value is based on the saturation value that was obtained earlier. Figure 6 shows (a) the thresholding image and (b) RGB on the thresholding (c) HSI image in the Thresholding image.

![Figure 5. Results of cropping the image](image)

2.3. Image training with artificial neural networks

The next stage is image training using artificial neural networks. At this stage the HSI value that has been successfully obtained in the image will be used as training input on artificial neural networks. The network architecture designed is 3 neurons in the input layer, 10 neurons in the hidden layer and 1 neuron in the output layer. The artificial neural network method used in this study is backpropagation with the sigmoid activation function[15–17]. Figure 7 shows the network training process with momentum 0.7 and performance goal 0.001. This network training uses the supervised learning method[18–20], so that from the beginning the network training output value has been determined first. In this network training the output value of good quality coffee beans = 1 and poor-quality coffee beans = 0.
2.4. Interface design
The next step is designing the application interface. At this stage we design the application interface for testing using MATLAB software. In figure 8 we can input the image of the coffee beans by using the 'load' button so that the image of the coffee beans will appear in the input image. Then the 'process' button will be processed in the input image so that the segmented image will be displayed in the box next to it. The results dialog box will display the results of detecting coffee beans whether good or bad.

3. Results and discussion
3.1. Training testing
The coffee bean samples used in this study were 100 raw coffee beans with the distribution of 40 good quality coffee beans and 40 poor quality coffee beans which had been separated to be used in the training process and in the testing process 10 good quality non-trained coffee beans were used and 10 poor quality non-trained coffee beans. The testing of the detection of the quality of coffee beans is divided into 2 parts:
1. The trial uses 30 training samples by dividing 15 good quality coffee beans and 15 poor quality coffee beans.
2. The trial uses 20 non-train samples by dividing 10 good-quality coffee beans and 10 poor-quality coffee beans. As shown in Table 1.

| Coffee Beans        | H     | S     | I     | Output |
|---------------------|-------|-------|-------|--------|
| Biji_Baik1          | 0.1170| 0.3604| 0.2376| 1      |
| Biji_Baik2          | 0.1187| 0.2973| 0.2524| 1      |
| Biji_Baik3          | 0.1166| 0.3664| 0.2367| 1      |
| Biji_Baik4          | 0.1049| 0.3831| 0.2279| 1      |
| Biji_Baik5          | 0.1119| 0.4024| 0.2336| 1      |
| Biji_Baik6          | 0.1200| 0.3307| 0.2476| 1      |
| Biji_Baik7          | 0.1234| 0.3317| 0.2316| 1      |
| Biji_Baik8          | 0.1181| 0.3762| 0.2301| 1      |
| Biji_Baik9          | 0.1195| 0.3760| 0.3019| 0.996 |
| Biji_Baik10         | 0.1207| 0.3025| 0.2085| 1      |
| Biji_Buruk11        | 0.0956| 0.4726| 0.1774| 0.027 |
| Biji_Buruk12        | 0.1174| 0.5604| 0.2462| 0      |
| Biji_Buruk13        | 0.0835| 0.5374| 0.1902| 0      |
| Biji_Buruk14        | 0.1102| 0.5782| 0.2647| 0      |
| Biji_Buruk15        | 0.1145| 0.5815| 0.2882| 0      |
| Biji_Buruk16        | 0.1154| 0.5435| 0.2890| 0      |
| Biji_Buruk17        | 0.1040| 0.5612| 0.2238| 0.0612 |
| Biji_Buruk18        | 0.0995| 0.5887| 0.1961| 0      |
| Biji_Buruk19        | 0.1055| 0.5500| 0.2608| 0      |
| Biji_Buruk20        | 0.1026| 0.5599| 0.2493| 0      |

Based on Table 1, it can be explained that the algorithm that has been designed is able to detect non-training data properly. This might be due to not too much non-training data. With more non-training data being tested, a detection error might be seen.

3.2. Application interface testing
The interface that has been designed successfully runs properly. Figure 9 shows the interface that was designed during the process of detecting the quality of coffee beans. After the image is successfully inputted, the next step is processing the image with the 'PROCESS' button. The results will display a segmented HSI image along with the results of quality detection in the image of the coffee bean.

Figure 9. Application interface testing

4. Conclusion
Based on the discussion above, the conclusions are as follows:
1. The quality detection system of coffee beans based on quality can be created based on differences in the intensity of the HSI value (Hue Saturation Intensity) in the image of coffee beans with digital image processing and backpropagation artificial neural networks.

2. Training of artificial neural networks designed in this study using optimal network architecture with Performance goals 0.001, Momentum 0.7.

3. For testing data that is not trained the system is able to detect quite well.

Acknowledgments
The authors are grateful to the Universitas Sumatra Utara for supporting this research. We also thank all the lecturers and staff of the Computer Laboratory and the Arrangement of the Department of Electrical Engineering at the Universitas Sumatra Utara who have helped this research.

References
[1] Wang N and Lim L T 2012 Fourier transform infrared and physicochemical analyses of roasted coffee J. Agric. Food Chem. 60 5446–53
[2] Nasution T H, Putramas A, Soeharwinto, Fahmi and Siregar I 2018 Automatic coffee roaster design using Arduino AIP Conference Proceedings
[3] Kelly C and Scott J 2015 Microwave aquametry of roasting coffee beans Asia-Pacific Microwave Conference Proceedings, APMC vol 3
[4] Kleinwächter M, Bytof G and Selmar D 2014 Coffee Beans and Processing Coffee in Health and Disease Prevention
[5] Santos J R, Sarraguça M C, Rangel A O S S and Lopes J A 2012 Evaluation of green coffee beans quality using near infrared spectroscopy: A quantitative approach Food Chem.
[6] Arboleda E R, Fajardo A C and Medina R P 2018 An image processing technique for coffee beans identification 2018 IEEE International Conference on Innovative Research and Development, ICIRD 2018
[7] Nasution T H and Andayani U 2017 Recognition of roasted coffee bean levels using image processing and neural network IOP Conference Series: Materials Science and Engineering
[8] Indonesia S N 2008 Biji Kopi
[9] Rao R A A, Mashor M Y, Ahmad R B, Noor S S M and KOsman M 2010 Comparison of colour thresholding method using RGB and HSI information for Ziehl-Neelsen sputum slide images 10th International Conference on Information Sciences, Signal Processing and their Applications, ISSPA 2010
[10] Chien C L and Tseng D C 2011 Color image enhancement with exact HSI color model Int. J. Innov. Comput. Inf. Control
[11] Blotta E, Bouchet A, Ballarin V and Pastore J 2011 Enhancement of medical images in HSI color space Journal of Physics: Conference Series
[12] Rahmat R F, Chairunnisa T, Gunawan D, Pasha M F, Budiarto R 2019 Hand gestures recognition with improved skin color segmentation in human-computer interaction applications Journal of Theoretical and Applied Information Technology 97(3), pp. 727-739
[13] Mitchell H B and Mitchell H B 2010 Image Thresholding Image Fusion
[14] Russ J and Neal F 2015 Segmentation and Thresholding The Image Processing Handbook, Seventh Edition
[15] Rojas R and Rojas R 2011 The Backpropagation Algorithm Neural Networks
[16] Han J and Moraga C 2012 The influence of the sigmoid function parameters on the speed of backpropagation learning
[17] Andayani U, Nababan E B, Siregar B, Muchtar M A, Nasution T H and Siregar I 2017 Optimization backpropagation algorithm based on Nguyen-Widrom adaptive weight and adaptive learning rate 2017 4th International Conference on Industrial Engineering and Applications, ICIEA 2017
[18] Baştanlar Y and Özuysal M 2014 Introduction to machine learning Methods Mol. Biol.
[19] Chelly S M and Denis C 2016 Applying supervised learning MathWorks, Inc.
[20] Anon 2010 Supervised learning Springer Tracts in Advanced Robotics