Journal of Research in Artificial Intelligence for Systems and Applications (RAISA)
2025, Vol. 01, No. 01, E-ISSN : XXXX-XXXX

Original Research

Optimization of Coffee Bean Maturity Classification by Segmentation on Multispectral Images Using HSV and DBSCAN

Muhammad Nurudin Hidayat [1], Tio Dharmawan [2*], Muhamad Arief Hidayat[3], Nelly Oktavia Adiwijaya [4]

[1] Information System, Computer Science Faculty, University of Jember, Jember, Indonesia
[2,3,4] Informatics, Computer Science Faculty, University of Jember, Jember, Indonesia

Abstract

In the coffee industry, sorting the maturity level of coffee beans is still done conventionally. In an effort to get good quality coffee beans, automatic classification of the maturity level of coffee beans is needed. The data in this research is multispectral image data and still has a background, so the preprocessing process is the main focus in this research to improve the performance of segmentation analysis in identifying objects and background image data. In the image data of 15 types of channels, a combination of 3 channel variations is carried out by applying HSV transformation so that the image data is easily processed by a computer, then the image data will be clustered using DBSCAN to identify coffee bean objects. The results obtained, the best channel combination in segmentation is blue, azure and amber, namely with a final weight value of 611. The segmentation results in the image data preprocessing process resulted in 100% accuracy. Meanwhile, the performance of the model without the segmentation preprocessing stage resulted in an accuracy of 92%. In conclusion, the performance of the model will be more optimal if preprocessing is done, namely segmentation in separating object and background data.

Keywords

Coffee Bean Maturity, Multispectral Image, DBSCAN, HSV.

1. Introduction

Coffee is one of the most important of all global food commodities [1]. It is therefore important for the sector to strive for product quality, as it is a major aspect of consumer choice[2]. The most important stage in improving the quality of coffee beans is the picking and sorting of ripe coffee beans [3]. The sorting stage of coffee beans is currently still done manually, so it takes a lot of time and is subjective because it depends on the perspective of the individual who is sorting[2].

Controlling the quality of coffee beans by classifying them during the harvest stage is a fundamental factor in obtaining better coffee quality so as to increase market value. In this case, it is necessary to develop specialized technologies to help reduce the time consumption of the sorting process and the consistency of coffee bean maturity standards. There are various variables that can be taken into consideration in determining the maturity of coffee beans, one of which is quantitative color measurement [3].

Classifying the maturity level of coffee beans can be done using multispectral images. Research conducted on the maturity level of Arabica coffee beans applies the multispectral concept using 15 color channels in image data [2]. The multispectral imaging technology is considered good in generating

*Corresponding author: Tio Dharmawan

Email addresses:

email1 muhammadnurudin01@gmail.com (Muhammad Nurudin Hidayat), tio.pssi@unej.ac.id (Tio Dharmawan), arief.hidayat@unej.ac.id (Muhamad Arief Hidayat), nelly.oa@unej.ac.id (Author Nelly Oktavia Adiwijaya)

Received: 13 12 2024; **Accepted:** 28 12 2024; **Published:** 15 01 2025

gradations due to its ability to generate large amounts of data. This technique provides measurements about the spatial configuration of objects and their spectral characteristics by generating three dimensions: two dimensions correspond to the conventional vision space, while the third dimension corresponds to the spectral response [2]. Based from these research references, this research conducts a more in-depth preprocessing process in an effort to improve the performance of the analysis, namely segmentation.

In this research, the image data used has 15 types of channels, each coffee image still has a background. The segmentation technique is carried out at the preprocessing stage to obtain a coffee bean image without background from the coffee dataset [4]. This is done so that the model can maximize training focus on the dataset object. Segmentation limits to a certain range of values to avoid the entire area being. Furthermore, 15 types of channels will be combined in 3 channel variations to find the boundary points between the coffee bean object and the background of the coffee bean image. The boundary points will produce color segmentation between the object and the background. The segmentation results are then subjected to feature extraction based on histogram characteristics. Characteristic parameters on the histogram are mean, skewness, variant, kurtosis and entropy [5].

In an effort to improve the performance of the classification of the maturity level of coffee beans, a preprocessing process is carried out on image data that still has a background. Segmentation technique is a solution to identify objects and backgrounds so that the model can focus on analyzing object data. Image segmentation using a combination of channels from the HSV transformation and the DBSCAN algorithm, is able to detect objects and image backgrounds. So that the results obtained are expected to produce an automatic coffee bean maturity level classification model that is precise and fast.

2. Related Works

There are many studies on the classification of coffee bean maturity. Tamayo-Monsalve et al.[2] proposed the use of convolutional neural network architectures VGG16, VGG19, InceptionV3, DenseNet201, and Inception-ResNetV2 on multispectral data which were explored in different experiments to extract characteristics from spectral images of coffee fruits in different stages of ripening to determine which one achieves the best results. In that aim, 4 experiments were conducted, applying imbalance balancing, subsampling, oversampling and weighting techniques on the training data. Afterwards, the aforementioned Deep Learning models were trained with and without applying transfer learning (TL) on the pre-trained models on the popular ImageNet dataset.

Syahputra et al. [6] proposed color features that can represent the character of coffee fruit maturity by conducting computer simulations to extract and calculate statistical values of color histograms and color moment values of four groups of coffee fruit. The results of the study using 200 coffee images show that the statistical value of the color histogram better describes the character of coffee fruit maturity, compared to the color moment. The kurtosis value of the color histogram has a different value for each category of coffee fruit maturity: young coffee has a kurtosis value of 17.2-28.3, semi-ripe coffee 29.2-31.4, ripe coffee 32.7-83.5 and mature coffee more than 84.2.

Suyoto et al. [5] proposed the use of feature extraction as a characterization method based on the histogram features of the image. The histogram displays the grayscale probability values of the pixel values in the image. The values contained in the resulting plot can be calculated using a number of first-order parameters properties include mean, skewness, variance, kurtosis, and entropy.

3. Research Method

3.1 Data Understanding

Data understanding is the initial stage in the image data processing process which aims to understand the data to be processed and identify problems in the data. Activities carried out at this stage are knowing the description of the dataset, such as the length and width of the dimensions, the number of channels, and knowing the visualization of the image data.

The dataset used in this study was obtained from the https://zenodo.org/ database. The data used is a 15-channel coffee dataset. The data has 5 labels namely dry, mature, semimature, overripe and immature as in Table 1. The image data has dimensions of 224 x 224 and 15 types of colors. Each channel represents a different wavelength as attached in Table 2.

Table 1. Data description

No	Ripening Stage	Number of images
1	Dry	78
2	Mature	160
3	Semimature	160
4	Overripe	112
5	Immature	130

Table 2. Channel description

No	Channel	Wavelength
1	*Violet*	*410*
2	*Royal Blue*	*450*
3	*Blue*	*470*
4	*Azure*	*490*
5	*Cyan*	*505*
6	*Green*	*530*
7	*Lime*	*560*
8	*Yellow*	*590*
9	*Amber*	*600*
10	*Red-Orange*	*620*
11	*Red*	*630*
12	*Deep Red*	*650*
13	*Far Red*	*720*
14	*NIR*	*840*
15	*NIR*	*950*

3.2 Preprocessing

This stage is the process of preparing the dataset before the modeling stage. The purpose of this stage is to prepare the image for the next processing stage. The preprocessing stage makes it easier for image data to be processed and increases the accuracy of the final image processing results.

In the combination 3 channels stage, the activity carried out is to combine 3 channels from a combination of 15 channels. The purpose of this stage is to obtain a more diverse combination of image data information and make it easier for the algorithm to identify objects with the background.

Then at the next stage the activity carried out is to find a good channel combination based on the color difference of the object and background. A good channel combination is when the image has a significant color difference between the object and background. From the results of this process, the best channel combination will be used at the Convert to HSV stage. at this stage changing the RGB color model to HSV. This stage aims to make the object easier to identify. Here is the formula for converting from rgb to hsv color model [7].

$$MAX = max\ (R, G, B) \quad (1)$$

$$MIN = min\ (R, G, B) \quad (2)$$

$$\delta = (MAX - MIN) \quad (3)$$

$$H = \left\{60\left(\frac{G-B}{\delta}\right),\ MAX = R\ 60\left(\frac{B-R}{\delta}+2\right),\right. \quad (4)$$

$$MAX = G\ 60\left(\frac{R-G}{\delta}+4\right),\ MAX = B\ tidak\ terdefinisi,\ MAX = 0$$

$$H\ = H + 360, H < 0 \quad (5)$$

$$S = \left\{\frac{\delta}{Max},\ Max \neq 0\ 0,\ Max = 0\right. \quad (6)$$

$$V = Max \quad (7)$$

3.3 Segmentation

Segmentation is a process that divides an image based on one or more criteria common to each part. Proper segmentation can reduce memory usage and increase image processing speed by reducing redundant information [8]. There are various techniques in segmentation namely thresholding, edge based, region based and clustering. In this study the segmentation used is thresholding. Thresholding refers to pixels with values greater than the threshold value (thresh) as white and pixels with values less than or equal to the threshold value as black or vice versa [9].

At this stage is the black and white color transformation stage based on the object color label at the previous stage. The white color represents the object color label, other than the object color label will be given black. The purpose of this stage is to facilitate the color cluster stage in the next analysis..

Then in the next stage is the clustering stage based on the contour mask results using DBSCAN. DBSCAN is an algorithm that has the ability to generate a variety of cluster shapes. Clusters are identified based on dot density. An area with high dot density indicates a cluster while an area with low dot density indicates an outlier cluster [10]. DBSCAN requires 2 inputs including MinPts which is the minimum number of points in the cluster and Eps which is the minimum distance between points to form a Neighborhood. To calculate the distance between points, the Euclidean distance formula is used as follows [11].

$$d = \sqrt{\sum \quad (x_i - y_i)^2} \quad (8)$$

The purpose of this stage is to produce white clusters. The best cluster is the cluster with the most white color points.

The best result obtained in the previous stage represents an object. Furthermore, the results in the data need cropping to separate the object and background. At this stage, cropping the object in the best cluster is done, so that the object data of the coffee bean image without background is obtained.

The evaluation of the object cropping result is done based on the number of successfully cropped object data and the length-width dimension of the cropping object. The best object cropping performance is if the object cropping result successfully crops the object correctly. Therefore, the dimension of the cropping object result is neither too small nor too big.

3.4. Feature Extraction

Feature Extraction is selecting a set of features to reduce dimensionality [12]. Feature extraction in images is divided into 4 categories: geometric features, statistical features, texture features and color features [13]. This stage is a process to extract color matrix data from each channel. The purpose of this process is to reduce the color matrix data on each channel. This stage uses the mean, variance, skewness, kurtosis, and entropy values. The following are the formulas for mean, variance, skewness, kurtosis and entropy [5].

a. Mean

$$\mu = \sum_{n=0}^{N} f_n P(f_n) \tag{9}$$

b. Variance

$$\sigma^2 = \sum_{n=0}^{N} (f_n - \mu)^2 P(f_n) \tag{10}$$

c. Skewness

$$\alpha_3 = \frac{1}{\sigma^3} \sum_{n=0}^{N} (f_{n-\mu})^3 P f_n \tag{11}$$

d. Kurtosis

$$\alpha_4 = \frac{1}{\sigma^4} \sum_{n=0}^{N} (f_n - \mu)^3 P(f_n) - 3 \tag{12}$$

e. Entropy

$$H = \sum_{n=0}^{N} P(f_n).^2 log\ log\ P(f_n) \tag{13}$$

3.5. Modelling

Modeling is done by dividing the data with K-fold. The K value used in this study is 10. From the K value, one fold from each other fold is used as testing data, while the remaining folds are used as train data. The model used in this study is SVM with a linear kernel.

the results of modeling are evaluated to determine how well the model built in the classification using accuracy, recall, precision, and F1-score in the model evaluation process.

4. Results and Analysis

The process started by combining 15 channels into many groups with 1 group consisting of 3 channels. The process resulted in 455 types of combinations. Here in Fig. 1 is an example visualization of some of the resulting combinations

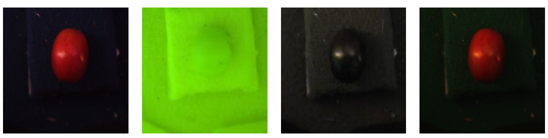

Figure 1. Sample image from combination

Then from the combination, 1 image from each class in each combination is taken. From the image, pixels of the object and background are taken, then the distance is calculated based on 6 types of basic colors. The closest distance of the 6 colors will be used as a label on the object and background. The following is an illustration of the color label retrieval in Fig. 2

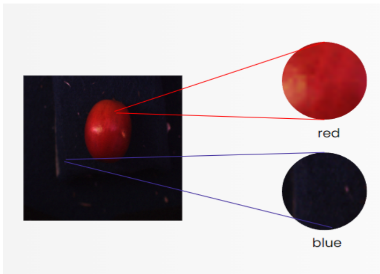

Figure 2. Sample colour from object and background

Table 3. Primary colour decimal code

No	Colour	Decimal Code RGB
1	*Red*	255,0,0
2	*Green*	0,255,0
3	*Blue*	0,0,255
4	*Cyan*	0,255,255
5	*Magenta*	255,0,255
6	*Yellow*	255,255,0

Then from the combination, 1 image from each class in each combination is taken. From the image, the pixels of the object and background are taken, then the distance is calculated based on 6 basic color types as in Table 3.. The closest distance from the 6 colors will be used as a label on the object and background. So that from this process 2275 data is formed. A combination is said to be good if the 5th class of the combination has a color label difference between the background and object. From this process, 68 types of combinations are obtained.

The next step is to convert from RGB color model to HSV color model. The results of the color label on the object in the previous stage will be used as a threshold. The color range in the HSV color model in OpenCV has a range from 0 to 180, so the 6 basic colors have a range of 30. Here is the color range for each basic color in Table 4. This process produces a binary image.

Table 4. . HSV Color range

No	Color	HSV decimal code	Lower Bound	Upper Bound
1	*Red*	0,255,255	0,0,0	15,255,255
2	*Red*	0,255,255	165,0,0	180,255,255
3	*Green*	60,255,255	45,0,0	75,255,255
4	*Blue*	120,255,255	105,0,0	135,255,255
5	*Cyan*	90,255,255	75,0 ,0	105,255,255
6	*Magenta*	150,255,255	135,0,0	165,255,255
7	*Yellow*	30,255,255	15,0,0	45,255,255

Figure 3. Mask of coffee bean image data

The mask result from the previous stage shown in Fig. 3 will be implemented using DBSCAN to cluster the coffee bean objects. The initial stage of this process is to store the coordinates of the white dots in an array variable, then the coordinate array data will be clustered with predetermined parameters. The coffee bean object is determined by the cluster that has the highest number of pixels as in Fig. 4. After that, the image will be cropped according to the cluster size. DBSCAN has 2 parameters, namely Eps and MinPts. In determining Eps, several experiments were conducted. The following in Fig. 5 is an evaluation of the Eps value.

Figure 4. Cluster segmentation result

From Fig. 5, the optimal eps is obtained with an optimal eps value of 2. Then at the next stage, an evaluation is carried out based on the length and width of the object in each maturity class. Furthermore, from the length and width of each object, the standard deviation value is calculated. Table 5 is an example of the standard deviation results of length and width at each maturity level.

Then from this value, weighting will be carried out. The weight value will be high if it has a low standard deviation, after which the weights of the standard deviation of length and standard deviation of width are summed up to produce a total weight on each class. The total weight is then summed up in the same combination in each class.

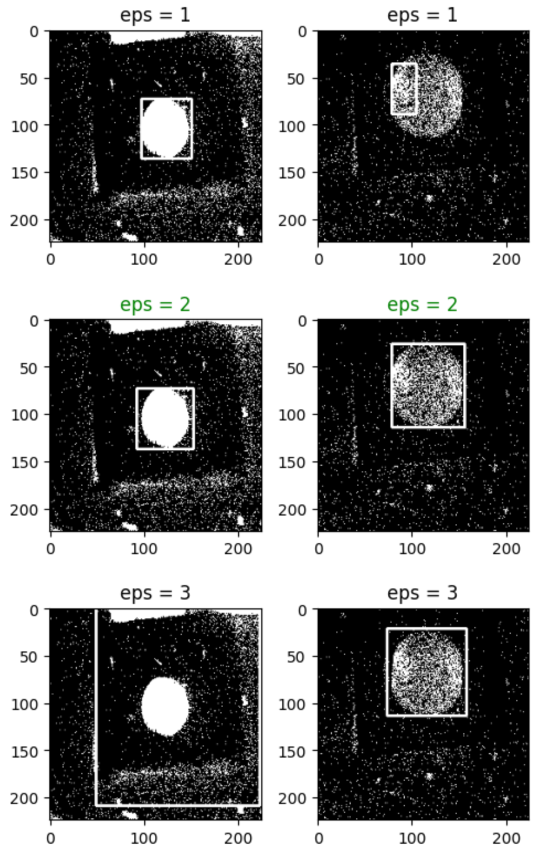

Figure 5. Evaluation of Eps value

The highest score is the best channel combination in segmentation. The best channel combination is blue, cyan and amber with a score of 611, from this channel combination as many as 638 images were successfully segmented by HSV and clustered by DBSCAN. The following in Table 6 is the result of the 5 highest scores from the channel evaluation

In the feature extraction stage, it is the process of extracting image data in the previous stage into table data using the mean, variance, kurtosis, skewness, and entropy values of the histogram value of each channel. The following in Fig 6 is an example of a histogram on channel red.

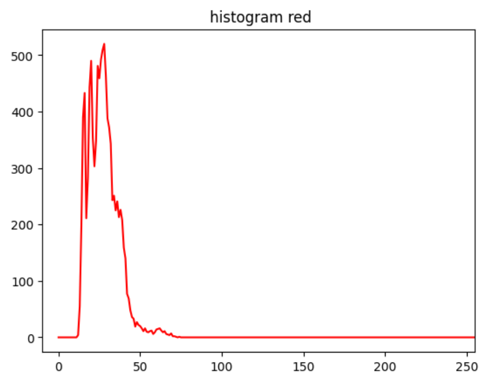

Figure 6. *Histogram channel red*

After the feature extraction stage, k-fold is then performed with a K value of 10. From this K value, one fold is used as testing data, while the remaining folds are used as train data.

Table 5.. Calculation of standard deviation of length and width

No	Class	combination	Std height	Std width	Score height	Score width	Total score
1	dry	violet, royal blue, amber	18.51	18.64	21	21	42
2		violet, royal blue, red	18.61	18.60	18	23	41
4		violet, royal blue, deep red	19.20	18.67	15	20	35
5	mature	violet, royal blue, amber	13.27	13.94932	30	31	61
6		violet, royal blue, red	14.04	14.59245	25	26	51
7		violet, royal blue, deep red	13.45	14.0611	29	28	57
8	semima-ture	violet, royal blue, amber	15.90	15.74976	39	38	77
9		violet, royal blue, red	21.16	21.06374	32	32	64
10		violet, royal blue, deep red	17.97	18.59667	35	34	69
11	overripe	violet, royal blue, amber	7.781105	7.319245	53	43	96
12		violet, royal blue, red	10.96646	12.87501	28	28	56
13		violet, royal blue, deep red	9.630879	9.818859	30	30	60
14	immature	violet, royal blue, amber	15.72354	20.06438	50	49	99
15		violet, royal blue, red	17.55575	17.55675	45	53	98
16		violet, royal blue, deep red	30.53554	31.48419	30	29	59

Table 6.. Channel evaluation score

No	combination	dry score	mature score	Semi Mature Score	Overripe Score	Immature Score	Total Score
1	blue, azure, amber	102	123	127	135	124	611
2	royal blue, blue, amber	96	118	136	124	121	595
3	royal blue, azure, amber	106	123	132	101	126	588
4	violet, blue, amber	128	106	126	84	130	574
5	blue, cyan, amber	100	98	126	124	97	545

The testing stage is divided into 2, namely by using segmentation and feature extraction, and without segmentation and feature extraction. The model used in the modeling process is Support Vector Machine using a linear kernel.

From the results of Table 7, it can be seen that testing using segmentation and feature extraction produces better accuracy than testing without segmentation and feature extraction. In testing using segmentation and feature extraction, the best accuracy is obtained by using an unbalanced scenario on a linear kernel type with a K-fold value of 10 resulting in an accuracy of 100%, while in testing without using segmentation and feature extraction obtained by using an unbalanced scenario on a linear kernel type with a K-fold value of 10 resulting in an accuracy of 98.75%.

Table 7. Evaluation score

No	Skenario	Acc	F1 Score	Recall	Precission
1	Without feature segmentation and feature extraction	98.75	98.25	98.14	98.6
2	With feature segmentation and feature extraction	100	100	100	100

5. Conclusion

The optimal channel combination for segmentation is blue, azure and amber with a final weight score of 611. From 640 image data, 638 image data were successfully segmented by HSV and clustered by DBSCAN. The classification performance results with segmentation and feature extraction on the classification of the maturity level of coffee beans produce 100% accuracy. While the performance results without using segmentation and feature extraction, resulted in an accuracy score of 98.75%. Classification optimization using segmentation can increase accuracy by 1.25%.

References

[1] W. R. Eustaquio, "Classification of Immature and Mature Coffee Beans Using RGB Values and Machine

Learning Algorithms," International Journal of Emerging Trends in Engineering Research, vol. 8, no. 7, Jul. 2020, doi: 10.30534/ijeter/2020/22872020.

[2] M. A. Tamayo-Monsalve et al., "Coffee Maturity Classification Using Convolutional Neural Networks and Transfer Learning," IEEE Access, vol. 10, pp. 42971–42982, 2022, doi: 10.1109/ACCESS.2022.3166515.

[3] S. Velásquez, A. P. Franco, N. Peña, J. C. Bohórquez, and N. Gutiérrez, "Classification of the maturity stage of coffee cherries using comparative feature and machine learning," Coffee Sci, vol. 16, 2021, doi: 10.25186/.v16i.1710.

[4] R. I. Hasan, S. M. Yusuf, M. S. Mohd Rahim, and L. Alzubaidi, "Automated masks generation for coffee and apple leaf infected with single or multiple diseases-based color analysis approaches," Inform Med Unlocked, vol. 28, Jan. 2022, doi: 10.1016/j.imu.2021.100837.

[5] R. Z. H. Suyoto, M. Komarudin, G. F. Nama, and T. Yulianti, "Classification of Civet and Canephora coffee using Support-Vector Machines (SVM) algorithm based on order-1 feature extraction," IOP Conf Ser Mater Sci Eng, vol. 1173, no. 1, p. 012006, Aug. 2021, doi: 10.1088/1757-899x/1173/1/012006.

[6] H. Syahputra, F. Arnia, and K. Munadi, "Karakterisasi Kematangan Buah Kopi Berdasarkan Warna Kulit Kopi Menggunakan Histogram dan Momen Warna," JURNAL NASIONAL TEKNIK ELEKTRO, vol. 8, no. 1, p. 42, Mar. 2019, doi: 10.25077/jnte.v8n1.615.2019.

[7] W. Cheni and Y. Q. Shii, "Identifying Computer Graphics using HSV Color Model and Statistical Moments of Characteristic Functions," International Congress on Mathematical Education, 2007.

[8] D. Oliva, M. A. Elaziz, and S. Hinojosa, "Studies in Computational Intelligence 825 Metaheuristic Algorithms for Image Segmentation: Theory and Applications." [Online]. Available: http://www.springer.com/series/7092

[9] Z. Niu and H. Li, "Research and analysis of threshold segmentation algorithms in image processing," in Journal of Physics: Conference Series, Institute of Physics Publishing, Jul. 2019. doi: 10.1088/1742-6596/1237/2/022122.

[10] S. Chakraborty and Prof. N.K.Nagwani, "Analysis and Study of Incremental DBSCAN Clustering Algorithm," International Journal of Enterprise Computing and Business Sytems, 2011, [Online]. Available: http://www.ijecbs.com

[11] E. Yohannes, F. Utaminingrum, and T. K. Shih, "Clustering of Human Hand on Depth Image Using DBSCAN Method," 2019. [Online]. Available: www.jitecs.ub.ac.id

[12] M. Suhaidi, R. Abdul Kadir, and S. Tiun, "A REVIEW OF FEATURE EXTRACTION METHODS ON MACHINE LEARNING," Journal of Information System and Technology Management, vol. 6, no. 22, pp. 51–59, Sep. 2021, doi: 10.35631/jistm.622005.

[13] W. K. Mutlag, S. K. Ali, Z. M. Aydam, and B. H. Taher, "Feature Extraction Methods: A Review," in Journal of Physics: Conference Series, IOP Publishing Ltd, Aug. 2020. doi: 10.1088/1742-6596/1591/1/012028.