

Classification of six banana ripeness levels based on statistical features on machine learning approach

Rifki Kosasih¹, Sudaryanto², Achmad Fahrurrozi¹

¹Department of Informatics, Faculty of Industrial Technology, Gunadarma University, Depok, Indonesia

²Departement of Industrial Engineering, Faculty of Industrial Technology, Gunadarma University, Depok, Indonesia

Article Info

Article history:

Received Jun 26, 2023

Revised Jul 15, 2023

Accepted Aug 18, 2023

Keywords:

Banana ripeness level

Features extraction

Histogram

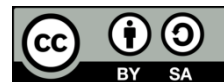
Naive Bayes classifier

Support vector machine

ABSTRACT

Banana plants are often cultivated because they have many benefits. In producing, we need to maintain the quality of bananas by looking at banana ripeness levels before being distributed to markets. The level of banana ripeness is related to marketing reach. If the marketing reach is far, bananas should be harvested when the ripeness level of bananas is still relatively low. A system that can classify the degree of ripeness of bananas can help overcome this problem. In this study, our dataset includes 6 ripeness levels of bananas, more than in previous related studies. Furthermore, we use the statistical features extraction method to find the parameters that affect the level of banana ripeness, considering the texture and color of the banana peel which determines the level of ripeness visually. The extraction used is features extraction based on a histogram, then we employ four features, i.e., mean, skewness, energy descriptor, and smoothness, generated from the image dataset. In the next stage, we perform classification based on the features that have been obtained. In this study, we use Naive Bayes classifier and support vector machine (SVM) algorithms. Based on the result of this research, the best performance is the Naive Bayes classifier, with an accuracy is 86.67%, a weighted average precision of 83.55%, and a weighted average recall of 86.67%.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Corresponding Author:

Rifki Kosasih

Department of Informatics, Faculty of Industrial Technology, Gunadarma University

Margonda Raya No. 100, Depok, 16424, Indonesia

Email: rifki_kosasih@staff.gunadarma.ac.id

1. INTRODUCTION

Banana is a fruit that contains vitamins, carbohydrates, and minerals. Banana plants grow in about 150 countries, i.e., India, Brazil, Philippines, Indonesia, China, Ecuador, Cameroon, Mexico, Colombia, and Costa Rica [1], [2]. Bananas are also an alternative fruit for tired people, so they are in great demand by many people. The high market demand, both for fresh consumption and for industrial raw materials, is an opportunity for agribusiness in Indonesia. One of Indonesia's most widely consumed bananas is the Cavendish because it has a delicious taste and soft texture.

Before shipping the harvested bananas, they are usually stored in a storage warehouse. Farmers need to know the level of banana ripeness during the storage process because it is closely related to the quality of the bananas. Therefore, farmers need to pay attention to the quality of bananas by knowing the banana's ripeness level. Thus, the farmer can know the right time and strategy for distributing bananas to markets.

One way to determine the level of ripeness of bananas is to look at the color change [3]. At first, the color changes in bananas can be known by looking with the human eye. However, this condition cannot be performed for a full day because it will cause the eyes to become tired, which results in an error in observing

color changes in bananas. Therefore, we need a system that can detect banana ripeness automatically. In this study, we propose to use a machine learning approach to classify banana ripeness levels. A machine learning algorithm is one part of artificial intelligence used to perform learning based on training data. Based on this learning, machine learning can perform classification [4], [5]. Several machine learning models can be used to classify, i.e., Naive Bayes classifier [6], [7] and support vector machine (SVM) [8], [9]. Several studies have been performed to detect the level of banana ripeness [10] using the SVM method to classify banana ripeness levels. His research used the input data of L^* , a^* , and b^* color features. The total data used is 210 images. Based on the study results obtained, an accuracy rate of 96.5%. The subsequent study was performed by [11] to classify a banana's ripeness level using the artificial neural network method based on the colour features of hue, saturation, value (HSV) and brown spots on bananas. From the study results, the level of accuracy obtained was 97.75%. However, the ripeness level is divided into 3 classes, i.e., not ripe, ripe and overripe.

Next, Tamatjita and Sihite [12] classify banana ripeness using HSV color space and the nearest centroid classifier. The data used are 60 banana images consisting of 4 classes, i.e., 15 green, 15 almost ripe, 15 ripe, and 15 overripe. Based on the research results obtained, an accuracy rate of 73.33%. Research from Kamelia *et al.* [13] uses the HSV and k -nearest neighbor color spaces to classify banana ripeness. The data used are 60 banana images divided into 30 training and 30 test data. However, the maturity level of bananas is divided into 3 classes: unripe, ripe, and rotten. The accuracy level obtained is 93.3%.

Furthermore, Wankhade and Hore [14] classify the ripeness of bananas using the SVM method based on feature extraction obtained from the inception v3 model. His research got an accuracy rate of 88.9%. However, it only uses 4 classes to determine the ripeness level, i.e., green, yellowish, mid-ripe, and overripe bananas. In another study, Anushya [15] determined the quality of bananas based on their level of maturity. His research classified banana ripeness using the decision tree and neural network method based on the gray-level co-occurrence matrix (GLCM) feature. The data used are 270 images, and the accuracy rate for the decision tree is 71%, while the neural network is 75.67%. However, the level of maturity used is only 3 classes, i.e., overripe, mid-ripe, and ripe. In other research, Prabha and Kumar [1] identified ripe bananas with color and size features obtained from banana images. Color features were obtained based on the average color intensity. At the same time, the size was calculated based on the characteristics of an area, the length of the main axis, and the length of the minor axis. In his research, two classifiers were used, one for color and the other for size features. The level of accuracy obtained is 99.1% for ripe bananas and 85% for unripe bananas.

Based on previous studies' descriptions, the banana's ripeness level is detected using color space features, i.e., L^* , a^* , and b^* [10] and HSV [11], or using feature extraction and machine learning algorithms. However, determining the class at the ripeness level is complex because the class definition uses the range of the color space. Therefore, in this study, we propose to detect the level of ripeness of bananas using other approaches, i.e., feature extraction of bananas texture using histogram-based texture applied on a greyscale image of bananas by converting its colour using a certain formula.

Furthermore, previous studies such as [13]–[15] classified the maturity level of bananas using 3 or 4 classes. The number of classes used in this study is still relatively small, so in this study, we also propose to use more banana ripeness levels than previous research, i.e., 6 classes. We also propose to compare two machine learning methods in classifying banana ripeness levels i.e., Naive Bayes classifier and SVM. Both algorithms are used because they provide good results based on previous studies. The resulting model is expected to classify banana ripeness better using machine learning algorithms with more level variations than previous studies.

2. RESEARCH METHOD

In this study, we construct two classification bananas ripeness models based on textural feature extraction, then compare two machine learning algorithms as classifiers for bananas' ripeness levels. The stages of this study as shown in Figure 1. The data collected were 45 images of bananas with 6 different ripeness levels. The data were randomly divided into 30 images as training data and 15 as test data.

Based on Figure 1, the pre-processing data stage includes converting image color and image rescaling step. In this stage, we pre-process input images by transforming a color image red green blue (RGB) into a grayscale image using (1).

$$I = 0.2989 * R + 0.5870 * G + 0.1140 * B \quad (1)$$

In the next step, we rescale the image size so that the image size becomes 150×200. After the pre-processing stage, feature extraction is carried out based on the texture of the banana images using the histogram method.

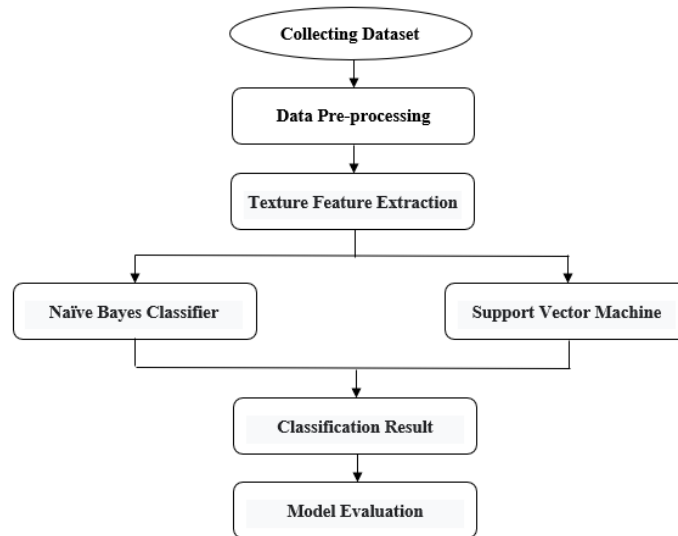


Figure 1. Stages of these research

2.1. Texture feature extraction

Features are particular characteristics of an object in the image that distinguish one object from another. The features that are commonly used are color features, shape features, and texture features. These features are then extracted to obtain the parameters used in the classification process. In color features, objects are usually recognized by the presence of color differences in a certain color space, i.e., HSV [16], YCbCr [17], RGB [18], and CIELAB color space [19]. Furthermore, in shape features, objects are usually recognized by the presence of differences in the shape of an object. One way to get these shape features is to use the Zernike moment [20]. Then, in texture features, objects are usually recognized by changing the object's texture. The texture is defined as a mutual relationship between the intensity values of neighboring pixels repeated in an area larger than the distance of the relationship. In this study, texture-based features were used, bearing in mind that the maturity level of bananas is determined based on the visual texture of the banana skin.

In this research, we use first-order statistical characteristics based on the characteristics of the histogram of the image, i.e., mean, skewness, energy, and smoothness [21], [22]. We calculate mean intensity using (2) based on the histogram results.

$$\mu = \sum_{j=0}^{M-1} j \cdot pr(j) \quad (2)$$

Where j is the grey level in the image f , $pr(j)$ represents the probability of occurrence j , and M represents the highest grey level value. The (2) produces the average brightness of the object. The second feature, we calculate skewness. Skewness is a measure of asymmetry for the mean intensity. The calculation formula is shown in (3).

$$skewness = \sum_{j=1}^{M-1} (j - \mu)^3 pr(j) \quad (3)$$

In skewness, a negative value indicates that the brightness distribution is skewed to the left for the mean, while a positive value indicates that it is skewed to the right for the mean. The third feature, we calculate energy descriptor. Energy descriptor is a measure that expresses pixel intensity distribution over the range of grey levels. The calculation formula is shown in (4).

$$E = \sum_{j=0}^{M-1} [pr(j)]^2 \quad (4)$$

A uniform image with one grey level value will have a maximum energy value of 1. The fourth feature, we calculate the smoothness of intensity in the image as shown in (5).

$$S = 1 - \frac{1}{1+\sigma^2} \quad (5)$$

In (5), if the S value is low, the image has a rough intensity. In the next step, we want to classify the ripeness level of bananas using a machine learning approach, i.e., Naive Bayes classifier and SVM.

2.2. Naive Bayes classifier

A Naive Bayes classifier is an algorithm-supervised classifier that classifies an object or data based on conditional probability [23]–[25]. The conditional probability calculates the probability that an event falls into one of several classes [26], [27]. Calculate the probability depending on the number of classes and the number of features. In this research, we have 6 classes (level of banana ripeness), i.e., $K_1, K_2, K_3, \dots, K_6$ with n features F_1, F_2, \dots, F_n . If given an event x , the formula for the probability of x entering one of the available classes is shown in (6).

$$\left. \begin{aligned} Pr(K_1|x) &= Pr(K_1) * Pr(F_1|K_1) * \dots * Pr(F_n|K_1) \\ Pr(K_2|x) &= Pr(K_2) * Pr(F_1|K_2) * \dots * Pr(F_n|K_2) \\ &\vdots \\ Pr(K_6|x) &= Pr(K_6) * Pr(F_1|K_6) * \dots * Pr(F_n|K_6) \end{aligned} \right\} \quad (6)$$

The conditional probability in (6) is obtained by using the probability density function (PDF) for the dataset in each class [25]. The result of this method is a vector whose components represent class predictions for each test data. We use the Naive Bayes classifier because this algorithm can recognize the different frequencies of the statistical features produced at the texture feature extraction stage using histograms. The more fluctuating the frequency of each statistical feature of each image in the dataset, the Naive Bayes classifier is expected to be able to classify it better.

2.3. Support vector machine

SVM is a supervised learning method to recognize patterns and classification [28]–[30]. In classification, SVM maps the input vector to a higher-dimensional space as shown in Figure 2. A hyperplane is constructed and used to separate the different classes in this space [31]. SVM will generate multiple hyperplanes, as shown in Figure 2(a). In the next stage, SVM will find the best maximum marginal hyperplane (MMH) in dividing the dataset into classes [32]. In SVM, there are support vectors, the data (points) closest to the hyperplane [33]. This point will define a hyperplane that separates sets of objects with different class memberships by calculating the margins [34]. Margin is the distance between the hyperplane and the closest class point. If the margin is more considerable between classes, it is considered a good margin. In other words, SVM will choose a hyperplane with the maximum margin, as shown in Figure 2(b). To select the optimal hyperplane, we use (7), (8), and (9) [35].

$$W^T x + b = 0 \quad (7)$$

In (7), W is a normal vector, which determines the direction of the hyperplane, and b is a displacement, which determines the distance between the hyperplane and the origin. Assume that the hyperplane can correctly classify the training data.

$$w^t x_i + b \geq 1, y = 1 \quad (8)$$

$$w^t x_i + b \leq -1, y = -1 \quad (9)$$

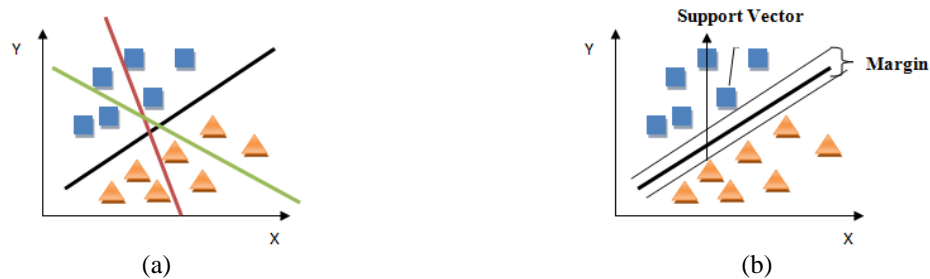


Figure 2. Classification by using SVM (a) SVM will generate multiple hyperplanes and (b) SVM will choose a hyperplane with the maximum margin

The maximum interval hypothesis is shown in (8) and (9). If the value of $y=1$, it indicates that the sample is positive, whereas if $y=-1$ is stated as a negative sample. Previous related studies about banana ripeness classification show that SVM gives promising results. In this study, SVM is used as a comparison to the Naive Bayes classifier, it also overcomes the problem of uniform frequency distribution that may appear in some features, which the Naive Bayes classifier cannot handle properly.

2.4. Model evaluation

In the next stage, we evaluate the model by calculating the value of precision, recall, and accuracy (Acc) in each category using (10), (11), (12), (13), and (14) [36]–[38].

$$Acc = \frac{TP+TN}{TP+FP+TN+FN} \times 100\% \quad (10)$$

$$p_i = \frac{TP_i}{TP_i+FP_i} \times 100\% \quad (11)$$

$$WAP = \frac{\sum_{i=1}^n p_i d_i}{\sum_{i=1}^n d_i} \quad (12)$$

$$r_i = \frac{TP_i}{TP_i+FN_i} \times 100\% \quad (13)$$

$$WAR = \frac{\sum_{i=1}^n r_i d_i}{\sum_{i=1}^n d_i} \quad (14)$$

Where WAP is weighted average precision, WAR is weighted average recall, p_i is the precision value for the i -th class, d_i is the number of actual data in the i -th class, and r_i is the recall value for the i -th class.

3. RESULTS AND DISCUSSION

In this section, we will discuss the research results obtained, i.e., the data collection results, the results from the classification of banana ripeness levels using the Naive Bayes classifier and SVM methods, and evaluate the classification model. The data collection results contain information regarding the amount of data per class and the appearance of some image data in the dataset. The classification results consist of information regarding the comparison of actual labels and predicted labels. The classification results are presented in the confusion matrix. The evaluation results include measuring model performance based on accuracy, recall, and precision.

3.1. Images dataset

This study shows banana images collected with 6 different ripeness levels in Figure 3. Our dataset is divided into two parts, i.e., 30 as training data and 15 as test data. Figure 3 shows 6 samples of bananas with 6 different ripeness levels. The first level is fresh green (Figure 3(a)), which is very good if harvested. The second is bright green (Figure 3(b)), ready for store pick up. The third is green with a hint of yellow, suitable for store stock (Figure 3(c)). The fourth is yellow with a touch of green (Figure 3(d)), ideal for store display and sale. The fifth is yellow and green-stemmed ready for consumption (Figure 3(e)). The sixth ripeness level is yellow with a few brown spots (Figure 3(f)), a banana with the best taste, smell, and nutrition.

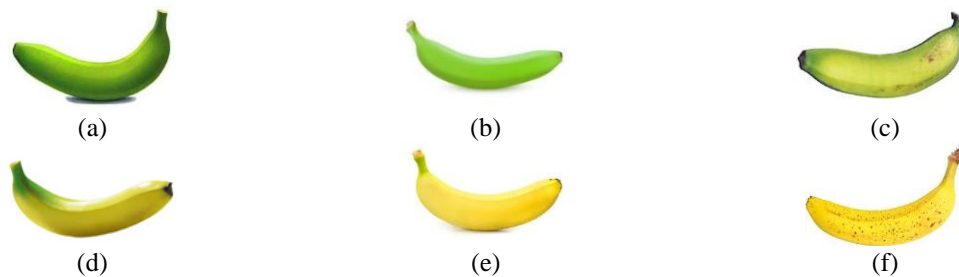


Figure 3. Sample of a dataset (6 of 45 banana images), (a) fresh green, (b) bright green, (c) green with a hint of yellow, (d) yellow with a touch of green, (e) yellow and green-stemmed, and (f) yellow with a few brown spots

3.2. Result of texture feature extraction

In this section, we will discuss feature extraction results. We perform feature extraction using the texture of the banana image. We propose to perform texture feature extraction because the formula used to obtain features is easier than feature extraction using color space. Besides that, determining the class at the ripeness level is complex because the class definition uses the range of the color space.

In this study, we use texture feature extraction based on the histogram, and there are 4 features extracted i.e., mean, skewness, energy, and image smoothness. In the next stage, we randomly divide the features into two parts, i.e., 30 data as training features and 15 as test features, as in Tables 1 and 2. As shown in Table 1, 30 data were used as training datasets (historical data). Each data has four features, i.e., mean, skewness, energy, and smoothness. Based on the extraction results, features with the same class have almost the same feature values, especially in the mean feature. Meanwhile, Table 2 shows that 15 banana images were used as a testing dataset, each having four features. The next step is to classify the test features using a machine-learning approach.

Table 1. Training features (20 of 30 training features)

No.	Mean	Skewness	Energy	Smoothness	Category
1	197.12700	-5.98550	0.27598	0.08603	1
2	208.51563	-2.93546	0.25212	0.05683	1
3	227.95467	-2.26523	0.36086	0.02915	2
4	248.32287	-0.33594	0.53966	0.00456	5
5	232.22179	-0.78187	0.26323	0.01682	3
6	231.47127	-0.79116	0.27344	0.01746	3
7	244.03913	-0.51267	0.40881	0.00781	4
8	222.00000	-8.37811	0.36913	0.05901	6
9	220.00000	-1.27179	0.35407	0.07378	6
10	220.30570	-2.75190	0.33190	0.04010	6
11	243.83170	-1.09295	0.27344	0.00960	4
12	199.75113	-3.94456	0.03376	0.06100	1
13	226.36540	-2.97633	0.21543	0.03164	2
14	234.04603	-0.79291	0.28622	0.01605	3
15	243.60570	-0.31211	0.37358	0.00623	4
16	248.47250	-0.31211	0.40285	0.00390	5
17	246.12530	-0.31770	0.52654	0.00570	4
18	246.11760	-0.19061	0.29540	0.00453	4
19	248.23440	-0.99380	0.49357	0.00770	5
20	249.86940	-0.22230	0.43305	0.00316	5

Table 2. Testing features

No.	Mean	Skewness	Energy	Smoothness
1	221.72137	-6.02257	0.23557	0.04961
2	214.56263	-2.10817	0.20874	0.03959
3	248.15070	-0.48039	0.34777	0.00502
4	244.16457	-0.07069	0.23265	0.00275
5	242.68340	-0.82140	0.21820	0.00890
6	244.60807	-0.70983	0.53331	0.00864
7	246.94830	-0.24114	0.23288	0.00367
8	225.37803	-2.51085	0.45057	0.03786
9	233.32313	-0.80087	0.17995	0.01639
10	237.19450	-1.64955	0.10263	0.01630
11	248.47250	-0.31211	0.40285	0.00390
12	209.26930	-4.31712	0.11563	0.05634
13	190.18540	-1.37339	0.01843	0.04834
14	234.04780	-0.79326	0.28792	0.01605
15	165.55020	-1.61060	0.00930	0.03150

3.3. Classification result using Naive Bayes and support vector machine

In this study, we want to compare two machine learning approaches, i.e., Naive Bayes classifier and SVM. Both algorithms are used because they provide good results based on previous studies. This section discusses the classification results using the Naive Bayes and SVM, as shown in Table 3. Based on Table 3, there are 15 test data selected randomly. We test using the Naive Bayes classifier and SVM methods. In the next step, we create a confusion matrix based on the prediction results in Table 3.

In Table 4, 4 bananas on actual condition in class 1 are predicted to enter class 1. In the second row, there is 1 banana on a real need in class 2 that is expected to join class 2. In the third row, there are 3 bananas

in the actual condition in class 3 predicted to enter class 3. In the fourth row, 3 bananas in the actual situation in class 4 are expected to join class 4. However, there is 1 banana on existing conditions in class 4 that are predicted to enter classes other than 4. In the fifth row, 2 bananas in the actual condition in class 5 are indicated to enter class 5. In the last row, on actual condition, there is 1 banana in class 6, but it is predicted to join class other than 6.

In Table 5, 3 bananas in the actual condition in class 1 are predicted to enter class 1. However, there is 1 banana on a real need in class 1 that is expected in other than class 1. In the second row, there is 1 banana on actual condition in class 2 that is predicted to enter class 2. In the third row, 2 bananas in the actual condition in class 3 are expected to join class 3. However, there is 1 banana in the actual situation in class 3 that is predicted in other than class 3. In the fourth row, 3 bananas in the actual condition in class 4 are expected to enter class 4. However, there is 1 banana in actual conditions. Bananas in class 4 are predicted to enter classes other than 4. In the fifth row, 2 bananas in the actual condition in class 5 are indicated to enter class 5. In the last row, in actual condition, 1 banana in class 6 is expected to join class 6.

In the next step, we evaluate by calculating accuracy, precision, and recall for the Naive Bayes classifier and SVM using (10), (11), (12), (13), and (14). The overall evaluation results are shown in Table 6. Based on Table 6, the model was constructed using the Naive Bayes classifier algorithm provided a WAP value of 83.6% and a WAR value of 86.7%. In contrast, we obtain a WAP value of 85.6% and a WAR value of 80% by using SVM.

Based on these results, the model built using the histogram and Naive Bayes classifier performs better than that of the histogram and SVM. These results showed that the Naive Bayes model performs well despite the small data. In addition, the Naive Bayes model can perform well in classifying banana ripeness with more variations in ripeness levels (6 classes) compared to previous studies, using a maximum of 4 classes.

Table 3. Classification result using Naive Bayes and SVM

No.	Actual	Naive Bayes	SVM
1	6	1	6
2	1	1	2
3	5	5	5
4	4	4	4
5	4	4	4
6	4	4	5
7	4	5	4
8	2	2	2
9	3	3	3
10	3	3	4
11	5	5	5
12	1	1	1
13	1	1	1
14	3	3	3
15	1	1	1

Table 4. Confusion matrix of Naive Bayes classifier

		Prediction					
		1	2	3	4	5	6
Actual	1	4	0	0	0	0	0
	2	0	1	0	0	0	0
	3	0	0	3	0	0	0
	4	0	0	0	3	1	0
	5	0	0	0	0	2	0
	6	1	0	0	0	0	0

Table 5. Confusion matrix of SVM

		Prediction					
		1	2	3	4	5	6
Actual	1	3	1	0	0	0	0
	2	0	1	0	0	0	0
	3	0	0	2	1	0	0
	4	0	0	0	3	1	0
	5	0	0	0	0	2	0
	6	0	0	0	0	0	1

Table 6. Overall evaluation result using Naive Bayes and SVM

Class	Naive Bayes classifier			SVM		Number of test data	
	Acc	Precision	Recall	Acc	Precision		Recall
1		80 %	75%		100%	75%	4
2		100%	100%		50%	100%	1
3		100%	100%		100%	66.7%	3
4		100%	100%		75%	75%	4
5		66.7%	100%		66.7%	100%	2
6		0%	0%		100%	100%	1
Weighted average	86.67%			80%			
		83.6%	86.7%		85.6%	80%	

4. CONCLUSION




The ripeness of bananas can usually be seen from the color change. Therefore, several studies have used the color space feature to detect bananas' ripeness levels. However, defining the range of values in the color space is complex, so another approach is needed to see the ripeness of bananas. In this study, we use the texture feature extraction method based on histograms to determine the parameters that affect the maturity level of bananas. This texture feature extraction produces several parameters, such as the average intensity, skewness, energy descriptor, and the smoothness of the intensity in the image. Based on the features that have been obtained, we classify the level of banana ripeness using the Naive Bayes method and SVM. The experiment results show that the best classification is the Naive Bayes classifier, with an accuracy of 86.67%, a WAP value of 83.56%, and a WAR value of 86.67%. For further research, we will add the amount of banana image data. In addition, another approach is used to detect the level of ripeness of bananas, i.e., using a deep learning algorithm, for example, a convolutional neural network, to detect the level of banana ripeness.

REFERENCES




- [1] D. S. Prabha and J. S. Kumar, "Assessment of banana fruit maturity by image processing technique," *Journal of Food Science and Technology*, vol. 52, no. 3, pp. 1316–1327, Mar. 2015, doi: 10.1007/s13197-013-1188-3.
- [2] R. Kosasih, "Classification of banana ripe level based on texture features and KNN algorithms," *Jurnal Nasional Teknik Elektro dan Teknologi Informasi*, vol. 10, no. 4, pp. 383–388, Nov. 2021, doi: 10.22146/jnteti.v10i4.462.
- [3] S. E. Adebayo, N. Hashim, K. Abdan, M. Hanafi, and M. Zude-Sasse, "Banana quality attribute prediction and ripeness classification using support vector machine," *ETP International Journal of Food Engineering*, 2017, doi: 10.18178/ijfe.3.1.42-47.
- [4] P. Vasavi, A. Punitha, and T. V. N. Rao, "Crop leaf disease detection and classification using machine learning and deep learning algorithms by visual symptoms: a review," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 12, no. 2, pp. 2079–2086, Apr. 2022, doi: 10.11591/ijece.v12i2.pp2079-2086.
- [5] Z. Mohammed, C. Hanae, and S. Larbi, "Comparative study on machine learning algorithms for early fire forest detection system using geodata.," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 10, no. 5, pp. 5507–5513, Oct. 2020, doi: 10.11591/ijece.v10i5.pp5507-5513.
- [6] W. Maharani and V. Effendy, "Big five personality prediction based in Indonesian tweets using machine learning methods," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 12, no. 2, pp. 1973–1981, Apr. 2022, doi: 10.11591/ijece.v12i2.pp1973-1981.
- [7] P. Shah, P. Swaminarayan, and M. Patel, "Sentiment analysis on film review in Gujarati language using machine learning," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 12, no. 1, pp. 1030–1039, Feb. 2022, doi: 10.11591/ijece.v12i1.pp1030-1039.
- [8] T. S. Lim, K. G. Tay, A. Huong, and X. Y. Lim, "Breast cancer diagnosis system using hybrid support vector machine-artificial neural network," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 11, no. 4, pp. 3059–3069, Aug. 2021, doi: 10.11591/ijece.v11i4.pp3059-3069.
- [9] M. J. Al Dujaili, A. Ebrahimi-Moghadam, and A. Fatlawi, "Speech emotion recognition based on SVM and KNN classifications fusion," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 11, no. 2, pp. 1259–1264, Apr. 2021, doi: 10.11591/ijece.v11i2.pp1259-1264.
- [10] J. C. Hou, Y. H. Hu, L. X. Hou, K. Q. Guo, and T. Satake, "Classification of ripening stages of bananas based on support vector machine," *International Journal of Agricultural and Biological Engineering*, vol. 8, no. 6, pp. 99–103, 2015, doi: 10.3965/j.ijabe.20150806.1275.
- [11] F. M. A. Mazen and A. A. Nashat, "Ripeness classification of bananas using an artificial neural network," *Arabian Journal for Science and Engineering*, vol. 44, no. 8, pp. 6901–6910, Aug. 2019, doi: 10.1007/s13369-018-03695-5.
- [12] E. N. Tamatjita and R. D. Sihite, "Banana ripeness classification using HSV colour space and nearest centroid classifier," *Information Engineering Express*, vol. 8, no. 1, p. 687, 2022, doi: 10.52731/iee.v8.i1.687.
- [13] L. Kamelia, M. R. Effendi, and N. H. Adila, "Ripeness level classification of banana fruit based on hue saturate value (HSV) color space using k-nearest neighbor algorithm," *International Journal of Advanced Trends in Computer Science and Engineering*, vol. 10, no. 2, pp. 941–947, Apr. 2021, doi: 10.30534/ijatcse/2021/651022021.
- [14] M. Wankhade and U. W. Hore, "Banana ripeness classification based on image processing with machine learning," *International Journal of Advanced Research in Science, Communication and Technology*, vol. 6, no. 2, pp. 1390–1398, Jun. 2021, doi: 10.48175/IJARSC-1571.
- [15] A. Anushya, "Quality recognition of banana images using classifiers," *International Journal of Computer Science and Mobile Computing*, vol. 9, no. 1, pp. 100–106, 2020.

- [16] T. Hidayat and A. R. Nilawati, "Identification of plant types by leaf textures based on the backpropagation neural network," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 8, no. 6, pp. 5389–5398, Dec. 2018, doi: 10.11591/ijece.v8i6.pp5389-5398.
- [17] Z. A. Alwan, H. M. Farhan, and S. Q. Mahdi, "Color image steganography in YCbCr space," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 10, no. 1, p. 202, Feb. 2020, doi: 10.11591/ijece.v10i1.pp202-209.
- [18] V. Jaiswal, V. Sharma, and S. Varma, "MMFO: modified moth flame optimization algorithm for region based RGB color image segmentation," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 10, no. 1, pp. 196–201, Feb. 2020, doi: 10.11591/ijece.v10i1.pp196-201.
- [19] O. P. Verma and N. Sharma, "Intensity preserving cast removal in color images using particle swarm optimization," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 7, no. 5, pp. 2581–2595, Oct. 2017, doi: 10.11591/ijece.v7i5.pp2581-2595.
- [20] G. Sucharitha and R. K. Senapati, "Shape based image retrieval using lower order Zernike moments," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 7, no. 3, pp. 1651–1660, Jun. 2017, doi: 10.11591/ijece.v7i3.pp1651-1660.
- [21] J. Deny and M. Sundararajan, "Survey of texture analysis using histogram in image processing," *International Journal of Applied Engineering Research*, vol. 9, no. 26, pp. 8737–8739, 2014.
- [22] A. Fahrurrozi and R. Kosasih, "Texture features and statistical features for wood types classification system," in *2022 5th International Conference of Computer and Informatics Engineering (IC2IE)*, Sep. 2022, pp. 186–191, doi: 10.1109/IC2IE56416.2022.9970071.
- [23] N. Hicham, S. Karim, and N. Habbat, "Customer sentiment analysis for Arabic social media using a novel ensemble machine learning approach," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 13, no. 4, pp. 4504–4515, Aug. 2023, doi: 10.11591/ijece.v13i4.pp4504-4515.
- [24] A. F. Sabbah and A. A. Hanani, "Self-admitted technical debt classification using natural language processing word embeddings," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 13, no. 2, pp. 2142–2155, Apr. 2023, doi: 10.11591/ijece.v13i2.pp2142-2155.
- [25] P. N. Andono, E. H. Rachmawanto, N. S. Herman, and K. Kondo, "Orchid types classification using supervised learning algorithm based on feature and color extraction," *Bulletin of Electrical Engineering and Informatics*, vol. 10, no. 5, pp. 2530–2538, Oct. 2021, doi: 10.11591/eei.v10i5.3118.
- [26] N. Kewsuwun and S. Kajornkasirat, "A sentiment analysis model of agritech startup on Facebook comments using Naive Bayes classifier," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 12, no. 3, pp. 2829–2838, Jun. 2022, doi: 10.11591/ijece.v12i3.pp2829-2838.
- [27] R. Kosasih and A. Alberto, "Sentiment analysis of game product on shopee using the TF-IDF method and naive bayes classifier," *ILKOM Jurnal Ilmiah*, vol. 13, no. 2, pp. 101–109, Aug. 2021, doi: 10.33096/ilkom.v13i2.721.101-109.
- [28] Y. Ben Salem and M. N. Abdelkrim, "Texture classification of fabric defects using machine learning," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 10, no. 4, pp. 4390–4399, Aug. 2020, doi: 10.11591/ijece.v10i4.pp4390-4399.
- [29] N. N. Moon *et al.*, "Natural language processing based advanced method of unnecessary video detection," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 11, no. 6, pp. 5411–5419, Dec. 2021, doi: 10.11591/ijece.v11i6.pp5411-5419.
- [30] I. M. Hayder, G. A. N. Al Ali, and H. A. Younis, "Predicting reaction based on customer's transaction using machine learning approaches," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 13, no. 1, pp. 1086–1096, Feb. 2023, doi: 10.11591/ijece.v13i1.pp1086-1096.
- [31] B. Abuhaija *et al.*, "A comprehensive study of machine learning for predicting cardiovascular disease using Weka and Statistical Package for Social Sciences tools," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 13, no. 2, pp. 1891–1902, Apr. 2023, doi: 10.11591/ijece.v13i2.pp1891-1902.
- [32] Y. K. Zamil, S. A. Ali, and M. A. Naser, "Spam image email filtering using K-NN and SVM," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 9, no. 1, p. 245, Feb. 2019, doi: 10.11591/ijece.v9i1.pp245-254.
- [33] S. Mishra, R. Kumar, S. K. Tiwari, and P. Ranjan, "Machine learning approaches in the diagnosis of infectious diseases: a review," *Bulletin of Electrical Engineering and Informatics*, vol. 11, no. 6, pp. 3509–3520, Dec. 2022, doi: 10.11591/eei.v11i6.4225.
- [34] A. Ambarwari, Q. J. Adrian, Y. Herdiyeni, and I. Hermadi, "Plant species identification based on leaf venation features using SVM," *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, vol. 18, no. 2, pp. 726–732, Apr. 2020, doi: 10.12928/telkomnika.v18i2.14062.
- [35] M. T. Ghazal and K. Abdullah, "Face recognition based on curvelets, invariant moments features and SVM," *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, vol. 18, no. 2, pp. 733–739, Apr. 2020, doi: 10.12928/telkomnika.v18i2.14106.
- [36] D. P. Lestari and R. Kosasih, "Comparison of two deep learning methods for detecting fire hotspots," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 12, no. 3, pp. 3118–3128, Jun. 2022, doi: 10.11591/ijece.v12i3.pp3118-3128.
- [37] C. Aroef, Y. Rivan, and Z. Rustam, "Comparing random forest and support vector machines for breast cancer classification," *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, vol. 18, no. 2, pp. 815–821, Apr. 2020, doi: 10.12928/telkomnika.v18i2.14785.
- [38] D. P. Lestari, R. Kosasih, T. Handhika, Murni, I. Sari, and A. Fahrurrozi, "Fire hotspots detection system on CCTV videos using you only look once (YOLO) method and tiny YOLO model for high buildings evacuation," in *2019 2nd International Conference of Computer and Informatics Engineering (IC2IE)*, Sep. 2019, pp. 87–92, doi: 10.1109/IC2IE47452.2019.8940842.




BIOGRAPHIES OF AUTHORS

Rifki Kosasih    received a B.S degree in mathematics from the University of Indonesia, Indonesia, in 2009, an M.S degree in mathematics from the University of Indonesia in 2012, and a Ph.D. degree from Gunadarma University, Indonesia, in 2015. He is a lecturer in the Department of Informatics (Gunadarma University). His research interests are image processing, object recognition, data science, manifold learning, machine learning, and deep learning. He can be contacted at email: rifki_kosasih@staff.gunadarma.ac.id.



Sudaryanto    received a B.S degree in industrial engineering from Bogor Agricultural University, in 1987, an M.S degree from the Asian Institute of Technology, Thailand, in 1994, and a Ph.D. degree from Aachen University of Technology, Aachen, Germany, in 2003. He is a lecturer in the Department of Industrial Engineering (Gunadarma University). His research interests are decision support systems, practical application of fuzzy sets theory, human factor (ergonomics), application of information and communication technology, supply chain management, and knowledge management. He can be contacted at email: sudaryanto@staff.gunadarma.ac.id.



Achmad Fahrurrozi    received a B.S. degree in mathematics from the University of Indonesia in July 2009, an M.S. in mathematics from the University of Indonesia in January 2012, and a Ph.D. degree from Gunadarma University - Indonesia, in 2016. His research interests include image processing, artificial intelligence, and statistical learning. He can be contacted at email: achmad.fahrurrozi12@gmail.com.