Journal of Mechanical, Civil and Industrial Engineering
ISSN: 2710-1436
DOI: 10.32996/jmcie
Journal Homepage: www.al-kindipublisher.com/index.php/jmcie

JMCIE
AL-KINDI CENTER FOR RESEARCH AND DEVELOPMENT

| RESEARCH ARTICLE

Identification of Banana Types with the Least-Squares Support Vector Machine (LS-SVM) Method

Michael Franch[1] ✉ Ronaldo Gozali[2], Jetly[3], Amir Mahmud Husein[4] and Christnatalis HS[5]
[12345]*Universitas Prima Indonesia, Indonesia*
Corresponding Author: Michael Franch, **E-mail**: michaelfranch07@gmail.com

| ABSTRACT
The classification of banana species is still done manually by banana farmers. This identification process has the disadvantage that it requires more manpower to sort, the level of perception of the type of fruit produced can be different because humans can experience fatigue, are not always consistent, and human judgments are also subjective. Thus, a tool is needed that can identify the type of banana fruit precisely and accurately. One of them is by creating a computer-based system using the statistical feature extraction method of digital images. By performing color feature extraction using Color Moments (RGBHSVYCbCr), then texture extraction using Gray-Level Co-occurence Matrix (GLCM), and using the Least-Squares Support Vector Machine (LS-SVM) method for classification of banana species. LS-SVM is a modification of SVM, which is used to improve classification performance. In the SVM algorithm, there is quadratic programming that is used to obtain the optimal solution in determining the Lagrange function; from the Lagrange function, it will be used in calculating the value of the weight and bias parameters. Quadratic programming is not efficient when applied to higher spatial dimensions because the computation will be very complex and very long. LS-SVM is better than standard SVM in terms of the calculation process, faster convergence, and higher precision. At the end of the experiment, the LS-SVM method succeeded in detecting the type of banana with a test accuracy value of 90%.

| KEYWORDS

Classification, Banana Species Detection, LS-SVM, GLCM, Color Moments

| ARTICLE INFORMATION

ACCEPTED: 26 August 2022 **PUBLISHED:** 02 September 2022 **DOI:** 10.32996/jmcie.2022.3.3.2

1. Introduction

Banana is a commodity that contributes significantly to national fruit production figures. Indonesian bananas supply the needs of not only the domestic market but also the international market. Therefore, the quality of bananas must be maintained [Dwi, 2017]. In bananas, characteristics such as size and color are used to classify. The classification of banana species is still done manually by banana farmers. This identification process has the disadvantage that it requires more human resources to sort, the level of perception of the type of fruit produced can be different because humans can experience fatigue, are not always consistent, and human judgments are also subjective. Thus, a tool is needed to identify the type of banana fruit precisely and accurately. One is creating a computer-based system using the digital image statistical feature extraction method [Nina, 2018]. Based on the results of previous studies using the GLCM, HOG, and SVM methods for the classification of banana species with an accuracy rate of 74.29% [Herry, 2020], GLCM to classify the quality of tangerines with an accuracy rate of 82.5% [Restu et al., 2018], HSV Color Moment and Local Binary with Naïve Bayes Classifier for types of food with an accuracy rate of LBP and HSV of 65%, where HSV is 65%, and LBP is 60%.

From existing research, the researchers used several methods, namely Color Moments (RGBHSVYCbCr) color feature extraction, texture extraction with Gray-Level Co-occurrence Matrix (GLCM), and Least-Squares Support Vector Machine (LS-SVM) classification speciesBanana fruit.

Based on the background that has been described, the researchers carried out a study entitled "Identification of Banana Types with the Least-Squares Support Vector Machine (LS-)VM) method." The purpose of this study is to provide convenience to the public in knowing the type of banana and be a reference based on the Utilization of Gray Level Co-occurrence Matrix (GLCM) and Color Moment using Least-Square Support Vector Machine.

2. Methodology

In this study, the research process carried out is as follows:

a. Data Collection
All data taken to carry out this research process is a banana dataset taken from the kangle website with several datasets taken, as many as 50 datasets.

b. Data Preprocessing
Before testing, the classification system is carried out with the Data Preprocessing process. This classification program will first enter each banana dataset consisting of 50 image datasets, including Ambon bananas, Barangan bananas, Kepokbananas, Mas bananas, and Red bananas.

c. Data Testing
This test was conducted to determine the accuracy of the LS-SVM method in identifying banana species.

Figure 1 below illustrates the proposed method for classifying banana species.

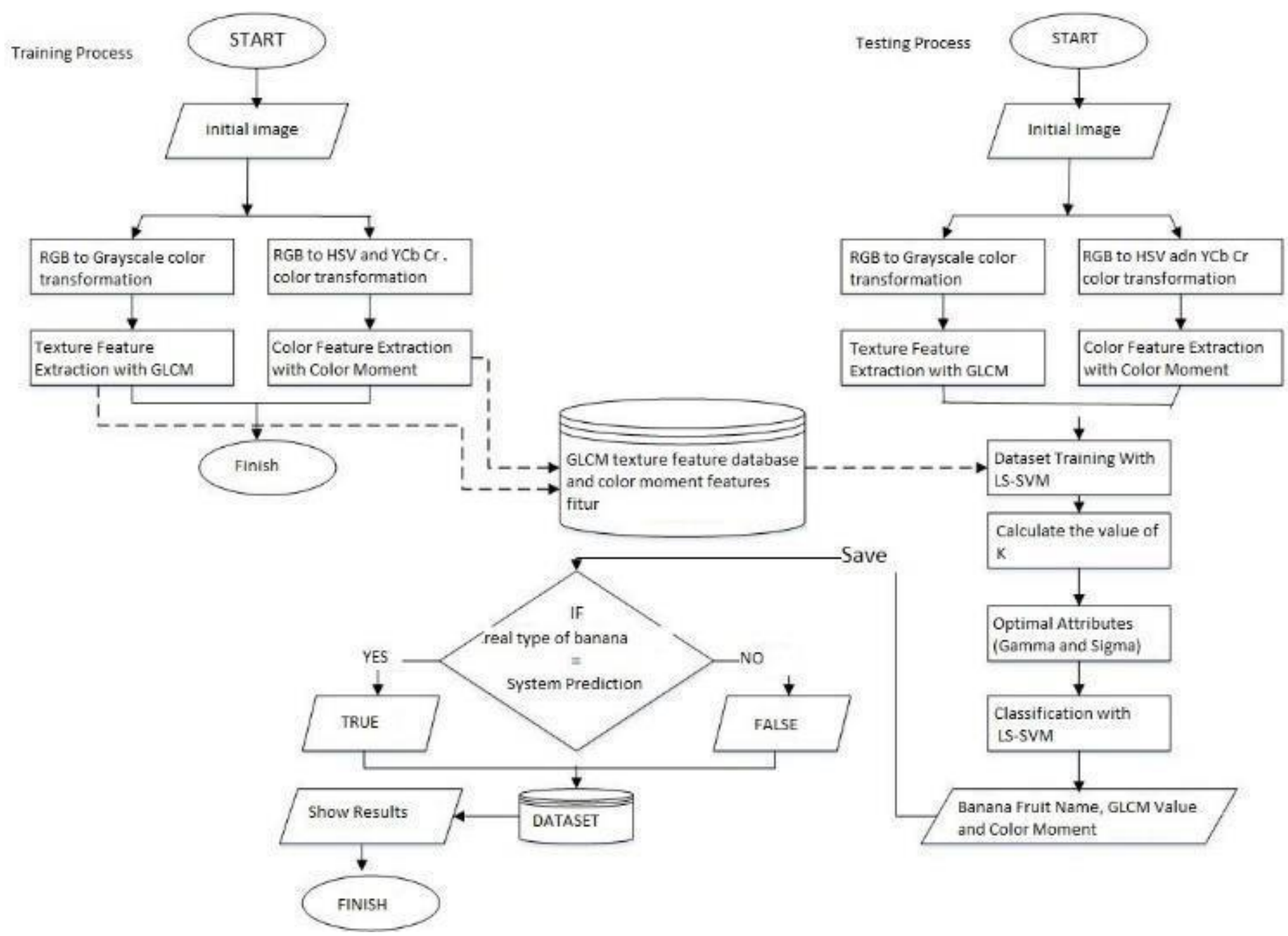


Figure 1. Process Analysis Flowchart

In Figure 1, the process begins with inputting the initial image, namely the RGB image, into the system. Image inputted will be transformed to HSV color and YCbCr color. The results of the RGB transformation are feature extraction using GLCM (Grey Level Co-occurrence Matrix). However, before performing feature extraction using GLCM, first RGB is transformed to Grayscale, and the results from Grayscale are used for GLCM feature extraction. Texture feature extraction is carried out to take the characteristics of the image, and this feature is very influential in the classification to determine the type of banana fruit. The steps in performing GLCM feature extraction (Grey Level Co-occurrence Matrix) are modeled in the flowchart in Figure 2.

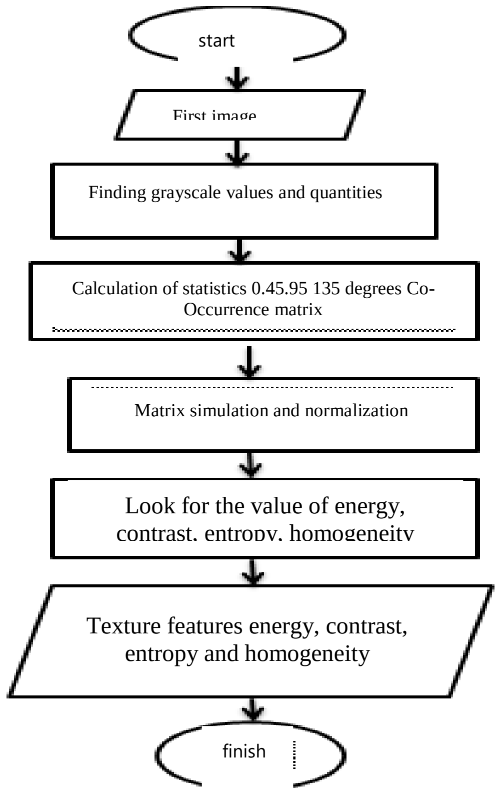


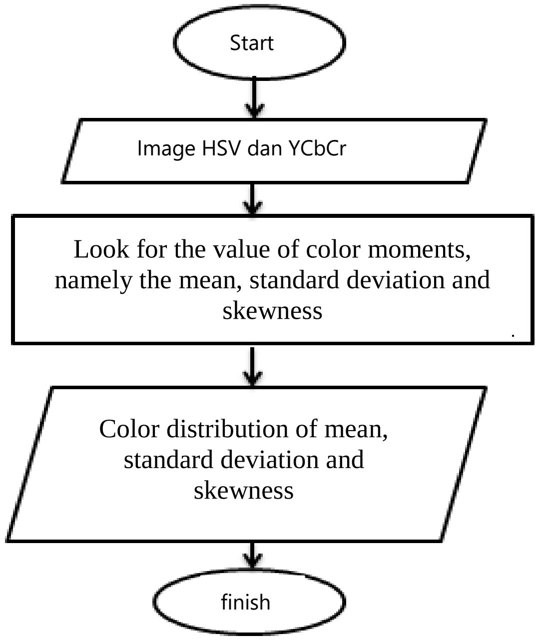
Figure 2. GLCM Feature Extraction Flowchart

An explanation of the process modeled in the flowchart in Figure 6 regarding the GLCM (Grey Level Co-occurrence Matrix) can be described as follows:

1. A color image is entered into the system.
2. The calculation process grayscale is used to add up all the R, G, and B values, then divided into 3, so that the average value is obtained and the results are entered into a new matrix.
3. Then do the quantization of the image, namely the Grayscale, which initially has a gray degree range of 0-255 to a gray level range of 0-15.
4. After looking for the quantization value, then the search for the value of the orientation of the angle is 0°,45°, 90°, and 135°. The process is done by counting how many pairs of pixels with the same neighboring pixels. Then the number is entered into a matrix.
5. In the symmetric, a transpose matrix is made based on the new matrix from the calculation results of process 5. A new matrix and a transpose matrix are added.
6. Matrix symmetric, then normalized. After the normalization is complete, all the normalized matrices are added.
7. The next step is calculating the feature extraction of Energy, Contrast, Entropy, and Homogeneity.
8. The results of the feature calculation process used are the texture feature values of Energy, Contrast, Entropy, and Homogeneity.

Color feature extraction is carried out by transforming RGB to HSV and RGB to YcbCr to find the value matrix color momentsColor feature extraction is carried out to distinguish images based on color. This feature is very influential in determining the type of banana fruit. The basis of Color Moments lies in the distribution of colors in the image, which can be interpreted as a probability distribution. The steps in performing the Color Moments are modeled in the flowchart in Figure 7.

Figure 3. Flowchart of color moments



An explanation of the process modeled in the flowchart in Figure 3.6 extraction feature Color Moments can be explained as follows:

1. The color images entered into the system are HSV and YCbCr images, and the two images will be processed separately, and comparisons will be made.
2. Then carry out the calculation process to get the distribution value of Mean, Standard Deviation Dan Skewness For each pixel, and the result is entered into a new matrix.
3. The result of the Mean, Standard Deviation, and Skewness calculation is an image with Color Moment.
4. The results of texture and color feature extraction can be stored as training data in a dataset and testing data for the banana fruit classification process with LS-SVM.

Least Squares Support Vectors Machine (LS-SVM) is a modification of SVM. Suppose SVM is characterized by quadratic programming with a constraint in the form of inequalities. On the other hand, LS-SVM is formulated using a constraint that is only an equation. So that the LS-SVM solution is generated by solving the linear equation; this is certainly different from SVM, where the solution is generated by solving quadratic programming. Currently, LS-SVM is mainly done on the classification and estimation of functions.

The following are the steps for classifying with LS-SVM:

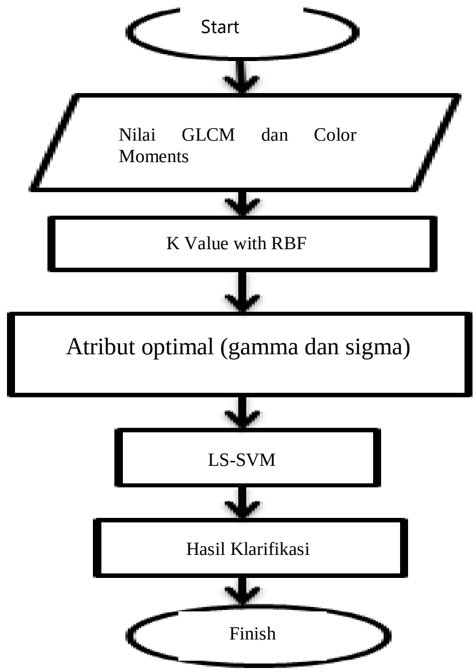


Figure 4. Flowchart of classification with LS-SVM

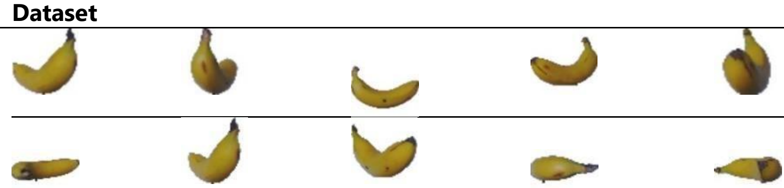
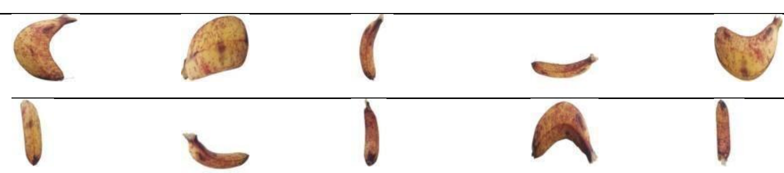
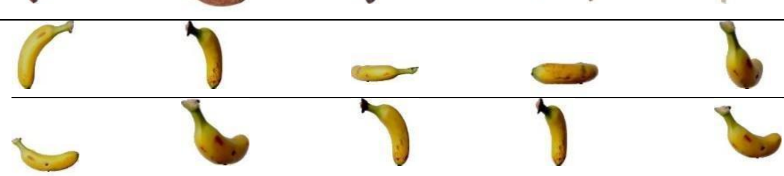
Analysis of the explanation of the process contained in Figure 8 regarding the classification of LS-SVM can be explained as follows:

1. Before classifying with LS-SVM, all the features in the Dataset must be trained first.
2. Process testing by calculating the value of K using the RBF kernel.
3. Entering testing so that gamma and sigma parameters are obtained for LS-SVM.
4. Entering gamma and sigma on LS-SVM are then processed and tested with testing.
5. If the process has been completed, it will get the classification results with LS-SVM.

4. Results and Discussion

In this experiment, we used five samples of bananas from 22 types in Indonesia. The bananas taken were Ambon bananas, Kepok bananas, Mas bananas, Barangan bananas, and Red bananas, % for training and 30% for testing.

Table 1 Dataset

Class	Dataset
Ambon	
Barangan	
Mas	
Kapok	

					
Merah					
					

And this is the final result for testing

identity	Ambon	Barangan	Mas	Kapok	Merah
	23.85%	19%	16.59%	21.12%	19.44%
	20.47%	18.37 %	20.32 %	20.03 %	20.81 %
	19.59 %	19.18 %	19.9 %	20.22 %	21.11 %
	18.31 %	19.66 %	22.27 %	19.03 %	20.72 %
	19.67 %	23.53 %	16.32 %	20.02 %	20.45 %
	20.04 %	20.82 %	20.7 %	18.32 %	20.12 %
	19.67 %	20.75 %	21.81 %	16.61 %	21.16 %
	20.45 %	21.25 %	21.27 %	16.53 %	20.51 %
	20.1 %	21.25 %	21.54 %	16.31 %	20.8 %
	20.66 %	18.42 %	20.29 %	20.2 %	20.43 %
	18.09 %	19.26 %	18.92 %	23.55 %	20.18 %
	19.68 %	18.3 %	20.09 %	21.31 %	20.62 %
	18.39%	21.25%	20.02%	16.67%	23,67 %
	20.7 %	21.12 %	20.71 %	16.94 %	20.53 %
	18.78 %	20.73 %	19.92 %	16.49 %	24.08 %
accuracy	19.90%	20.20%	20.05%	18.90%	20.95%

To perform the process of adding datasets to the banana type detection system, the user can click the Preprocessing menu so that the system will display the Training form, as shown in Figure 6.

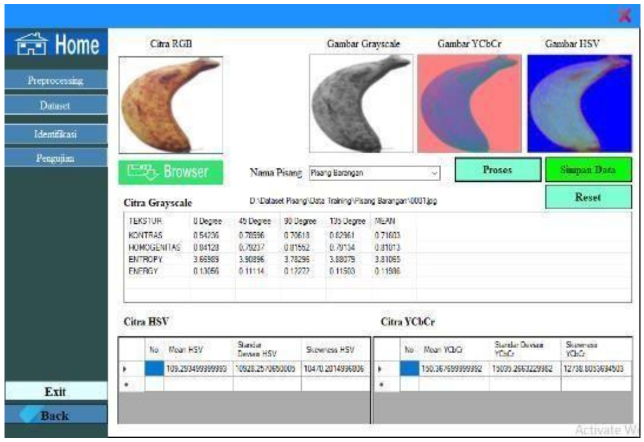


Figure 6. ***Train Forming***

Click the Save button to save the data entered into the *database*. To view *a database* that has been previously stored, then you can click the *Dataset*, so the system will display a database form, as shown in Figure 7.

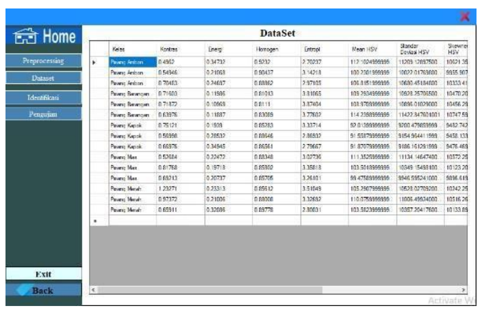


Figure 7. ***Dataset Form***

After that, the user can identify bananas by clicking the identification menu so that the system will display *the form* Identification Process.

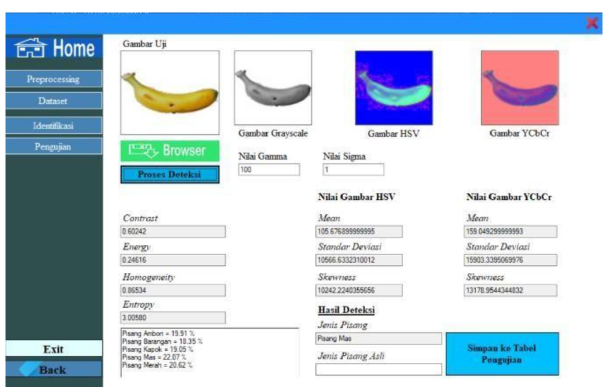


Figure 8. Identification Process Form

Based on the testing process carried out on the software, the following information can be obtained:

1. The Least-Squares Support Vector Machine (LS-SVM) method can detect banana species with high accuracy. Tall one.
2. The number of datasets will affect the accuracy of identifying banana species using the Least-Squares Support Vector Machine (LS-SVM) method. However, the more datasets, the execution process will take longer.
3. The test accuracy value obtained is 90%.

5. Conclusion

Finally, from the discussion in the previous chapters, the research in this final project can be drawn several conclusions, among others.

1. The Least-Squares Support Vector Machine (LS-SVM) method can detect and recognize various types of bananas with varying banana image input. The
2. accuracy of the banana species detection results with the Least-Squares Support Vector Machine (LS-SVM) method depends on the number of datasets contained in the system
3. The test accuracy value obtained is 90%

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers.

References

[1] Dwi Y (2017) Retno Nugroho Whidhiasih, Maimunah, Klasifikasi Tahap Kematangan Pisang Ambon Berdasarkan Warna Menggunakan Naive Bayes, *Jurnal Penelitian Ilmu Komputer, Sistem Embedded & Logic 5*(2) : 60-67.

[2] Herry K S (2020) Novan Wijaya, Klasifikasi Jenis Pisang Menggunakan Support Vector Machine dengan Fitur GLCM dan HOG, *Indonesian Journal of Computer ScienceISSN 2302-4364(print) dan 2549-7286(online), 9*(2).

[3] Nina S (2018) Jayanti Yusmah Sari, Ika Purwanti Ningrum Purnama, Identifikasi Tingkat Kematangan Buah Pisang Menggunakan Metode Ektraksi Ciri Statistik Pada Warna Kulit Buah, ISSN 2085-4552, ULTIMATICS, X(2).

[4] Nursuci P H, Nursanti N A, Putri N R, Agus Z A, and Darlis H (2017), LEAST SQUARES SUPPORT VECTOR MACHINES PARAMETER OPTIMIZATION BASED ON IMPROVED ANT COLONY ALGORITHM FOR HEPATITIS DIAGNOSIS, Jurnal Ilmu Komputer dan Informasi *(Journal of Computer Science and Information). 10(*1) 43-49,

[5] Restu W, Agus W W and Arry S (2018), Pemanfaatan Ciri Gray Level Co- Occurrence Matrix(GLCM) Citra Buah Jeruk Keprok (Citrus reticulata Blanco) untuk Klasifikasi Mutu, Jurnal Pengembangan *Teknologi Informasi dan Ilmu Komputer, 2* (11) hlm. 5769-5776, e-ISSN: 2548-964X, http://j-ptiik.ub.ac.id.

[6] Sugiartha G. R. A., Sudarma M dan Widyantara I. M. O, (2017) Ekstraksi Fitur Warna, Tekstur dan Bentuk untuk Clustered-Based Retrieval of Images (CLUE), *Teknologi Elektro, 16*(1).

[7] Yusuf E Y and Nur N (2021) Klasifikasi Jenis Pisang Berdasarkan Fitur Warna, Tekstur, Bentuk Citra Menggunakan SVM dan KNN, *Journal of Computer, Information System, & Technology Management, 4(*1) 28-36, Online ISSN:2615-7357, Print ISSN:2615-7233.

[8] Zainal A. dan Dyah S W. K. (2018) Penggunaan Analisa Faktoruntuk Klasifikasi Citra Penginderaan Jauh Multispektral, JUTI:*Jurnal Ilmiah Teknologi Informatika*, 1.

[9] Zeni D L (2019) Sistem Klasifikasi Jenis Pisang Berdasarkan Ciri Warna HSV Menggunakan Metode K-NN, Seminar Nasional Teknologi Informasi dan Komunikasi 2019, "Teknologi Humanis di Era Society 5.0", e-ISSN: 2685-5615.