Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 6, Number 2, April 2022
DOI : https://doi.org/10.33395/sinkron.v7i2.11333
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Identification of Buni Fruit Image Using Euclidean Distance Method

Dadang Iskandar Mulyana[1)], Abdul Hafidz[2)], Dyan Bagus Sumantri[3)], Kurniawan Setyo Nugroho[4)]
[1)2)3)4)] College of Computer Science Cipta Karya Informatics, Indonesia
[1]mahvin2012@gmail.com, [2]hafidabdul128@gmail.com, [3]bagusdyan12.db@gmail.com, [4]setyok61@gmail.com

Submitted : Mar 2, 2022 | **Accepted** : Mar 21, 2022 | **Published** : Apr 1 , 2022

Abstract: Identification is an important part of image analysis because in this procedure the desired image/image will be analyzed for further processing to make it easier to analyze for further purposes, for example in image identification pattern recognition which is part of image analysis used to divide an image into several parts and take some of the desired objects. This study aims to identify buni fruit with Euclidean distance and extract shape and texture features. Extraction of shape features using boundary metrics and whimsy. This boundary is considered to be able to recognize objects based on their shape and can distinguish them from other objects. For identification expositions, Euclidean distance is used which serves to represent the level of similarity of two images that take into account the distance value of the Euclidean distance. From the results of the evaluation using a disarray network by calculating precision, review, and accuracy, in order to identify the image of the buni fruit object properly.

Keywords: Identification; Euclidean Distance; Texture; Color; Buni Fruit Image.

INTRODUCTION

Buni in scientific language is called Antidesma bunius (L.) Spreng. This plant is a tree whose height can reach 15-30 m, trunk diameter of about 20-25 cm, many branches and shady. The male and female flowers of each buni are located on different trees, arranged in the form of panicles. The size of the female flower is larger than the male flower. The color of the buni fruit is initially bright green, when ripe it becomes red. Buni fruits are arranged in bunches, round or ovoid, about 3 cm in diameter (National Institute of Biology 1977).

According to Heyne (1987), small buni fruits are larger than peas, at first very sour red, then blackish and juicy with a sweet and sour taste. However, many people do not know about this plant. This type of buni plant can be identified based on the characteristics seen from its shape and texture. However, most people judge that the types of buni plants have a similar shape so it is difficult to distinguish them. So we need a system that can identify the image of medicinal plants of the type of buni based on the characteristics of their shape and texture. Image processing can be a solution in identifying images so that they can provide information about the buni fruit. Image processing is a field that studies how an image is formed, managed, and analyzed to obtain useful information. Image processing can produce information that can then be useful in helping humans work.

In image processing to identify an object, it is necessary to recognize the characteristics or features of the image to be identified, or what is usually called feature extraction. Feature extraction or feature extraction is the quantization of image characteristics into a group of appropriate feature values. Feature extraction which is usually used is the extraction of shape and texture features. The shape is one of the recognizable characteristics of an object to find the difference between the object and other objects. Meanwhile, texture feature extraction is a feature of an image that can be used to explore the characteristics of an image because the surface arrangement of the image contains information that can be utilized. After the object can be identified, then identification expositions can be done easily. Image identification is an exposition of recognizing images by grouping images on the image to obtain classes or categorize them based on certain characteristics. One of the boundaries that can be used for image identification is the Euclidean distance.

Euclidean distance is an image matching technique that can identify other images that are similar or have similarities. Euclidean distance represents the level of similarity of two images that take into account the distance

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 6, Number 2, April 2022
DOI : https://doi.org/10.33395/sinkron.v7i2.11333
e-ISSN : 2541-2019
p-ISSN : 2541-044X

value from Euclidean, if the smaller the Euclidean distance, the more similar the image. Euclidean distance is a calculation of the distance from two points in Euclidean space, which is used to study the relationship between angles and distances.

Expositions are carried out by comparing the proximity of the distance values of two variables, namely between the test image and the reference image to find the closest distance value. Several previous studies have shown that this method produces good accuracy in the application of image processing. Previous research related to the application of the Euclidean distance method for distance identification. In this study, the method used resulted in the success percentage reaching 87.84%. Another research is about identifying texture and color images using Euclidean distance. In this study, the verification of the success rate of the method used reached 84.00%. Subsequent research on the classification of body hydration levels based on urine color. By doing 20 experiments using the Euclidean distance method, the accuracy rate is 76%. This study aims to identify buni fruit using Euclidean distance and the shape and texture characteristics of the extract. Extraction of shape features using boundary and uncertainty metrics. This limit is considered to be able to recognize objects based on their shape and can distinguish them from other objects.

LITERATURE REVIEW

A similar study was conducted by N. Sivi Anisa and T. Herdian Andika (2020). Sivi Anisa and T. Herdian Andika (2020) conducted a study on a segmentation-based leaf image identification system using the K-Means clustering method. Sivi Anisa and T. Herdian Andika (2020) utilize image segmentation to select optimal feature images for analysis purposes.

Similar research has also been carried out by D. Nurnaningsih, D. Alamsyah, A. Herdiansah, and A. A. J. Sinlae (2021) with the title of image identification of rhizome medicinal plants with euclidean distance based on shape and texture characteristics. In the content of the research, there is feature extraction using metric and eccentricity parameters. This parameter is considered to be able to recognize objects based on their shape and can distinguish them from other objects. Meanwhile, texture feature extraction uses (GLCM) with parameters contrast, correlation, energy, and homogeneity.

Then also by CN Prabiantissa, AR Tri, and RA Asmara (2017) from the State Polytechnic University of Malang the title of the research is the identification system of natural batik and synthetic batik based on image color characteristics with the K-Means clustering method using the help of computerized camera media as a batik image capture. then calculate the RGB normalization value. Research conducted by C. N. Prabiantissa, A. R. Tri, and R. A. Asmara (2017) contains about automatically distinguishing natural and synthetic batik with the help of image processing applications. From the identification results obtained, it produces 2 outputs, namely natural batik and synthetic batik.

METHOD

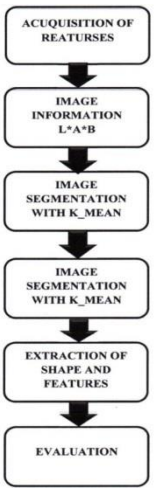


Figure 1. Research Stages

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 6, Number 2, April 2022
DOI : https://doi.org/10.33395/sinkron.v7i2.11333
e-ISSN : 2541-2019
p-ISSN : 2541-044X

For research carried out to run well, planned and structured, it is necessary to carry out the stages of research carried out. The stages of the research carried out include the process flow from the start of image acquisition then proceed to the L*A*B image transformation stage, then to the image segmentation stage with the K-Means method, to the extraction stage of shape and texture features and the final stage of evaluation can be seen in the following:

Image Acquisition

The initial stage is image acquisition, where images of medicinal plants of the Buni species are taken which are used as training information and test information. Expositions of image retrieval through direct independent shooting, where the image is photographed with the same light level. Information from the image acquisition used is 4 buni medicinal plants that are often found and easy to obtain, including Ginger, Galangal, Kencur and Turmeric. The distribution of the dataset uses an experimentation approach, which determines the structure of the model by dividing the information into half training and half testing. The number of datasets that have been collected is 80 images, with training information of 40 images and test information of 40 images.

Image Transformation L*a*b

The transformation using the L*a*b color space aims to identify advanced color content. There are 12 kinds of colors generated in the color wheel namely red, yellow, green, cyan, blue, and maroon, with all colors. The steps taken are to change and transform the image color space from RGB to XYZ. Furthermore, the results of the RGB color values are used as values to calculate the values of L*, a* and b*. This step is used to simplify segmentation expositions.

Image Segmentation with K-means Clustering

K-Means is an algorithm to group n objects based on attributes into k partitions, where k < n. The steps in the K-implies grouping algorithm begin with determining the number of clusters. To determine the centroid value, you can use the following formula

$$\overline{V_{ij}} = \frac{1}{N}\sum_{k=0}^{n_i} x_{kj}$$

Where vij is the centroid/average of the I-th bunch for the j-th variable. For Ni is the amount of information that is a member of the I-bunch. Then i,k is the index of the bunch j is the index of the variable. While xkj is the value of the k-th information in the bunch for the j-th variable.

Texture Feature Extraction

Feature extraction is a process to get the distinguishing features that distinguish an object from other objects. The extracted features are then used as boundaries or input values to distinguish objects from one another at the classification stage. In this study, the extraction of shape and texture features was used. To extract shape features, use boundary metrics and unpredictability. Metric is the ratio between the area and the circumference of an object. Meanwhile, unpredictability is the comparison value between the focal distance of the minor ellipse and the focus of the city hall leader ellipse of an object.

Image Identification with Euclidean Distance

Euclidean Distance is the matrix most often used to calculate the similarity of two vectors. Distance is the basis of the square contrast between 2 vectors. Two vector elements can measure each other by calculating the distance between them, or vice versa, to determine the degree of similarity. There are many distance estimates used in the sequence of visual examples.

Evaluation.

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 6, Number 2, April 2022
DOI : https://doi.org/10.33395/sinkron.v7i2.11333
e-ISSN : 2541-2019
p-ISSN : 2541-044X

For evaluation in this study using a disarray grid, where calculations will be made for accuracy, review, and precision. The test results will be entered into a disarray grid. Disarray grid consists of genuine positive, misleading positive, genuine negative and bogus negative to calculate precision, review and accuracy. Genuine positive is positive information that is predicted to be true. On the other hand, genuine negative is negative information that is predicted to be true. While genuine negative is negative information but it is predicted as positive information. In general precision, review and accuracy.

RESULT

In the development of a classification system for the level of maturity of pineapples, the training process is carried out first. The dataset collected is 80 images of pineapples. The image used in the training is 40 information images or half of all out images. Then, 40 training information. After the dataset is collected, the next step is to prepare for the training. Training and testing are carried out using the Matlab application. The initial stage in the process of transforming the color space from Red, Green, Blue images to L*a*b images. This is done so that the color content can be identified computerized. The compositions of the transformed images from RGB to L*a*b in the Matlab application can be seen in Figure 2 below.

Figure 2.(a) RGB image (b) Transformed Image L*A*B

After the image is converted to L*a*b, to make it easier to share the image, the image is converted or transformed into binary form. The output of these expositions is a binary image, where the desired object has a value of 1 (white color), while the foundation value is 0 (black color). The image from the transformation to binary is separated from the object with the foundation so that the obtained image can be used as a covering for further expositions. The resulting binary image is transformed into a grayscale image. It aims to simplify the image so that it is easy to expositions image processing. Figure 3. below is the result of the grayscale image transformation. Figure 3 below is the result of binary and grayscale image transformation.

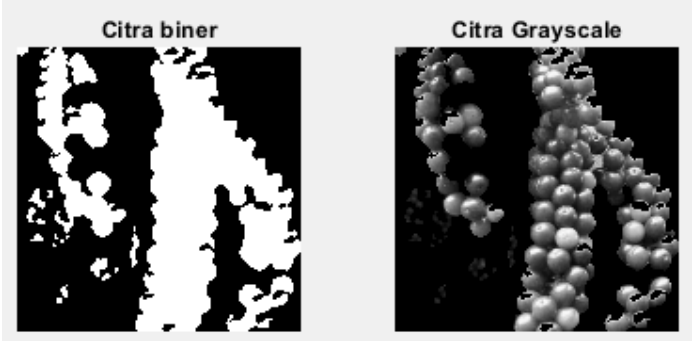


Figure 3. (a) Binary Transformation Results (b) Grayscale Image Transformation Results

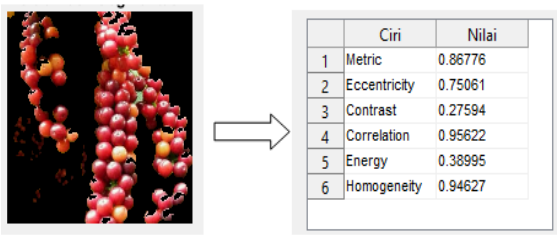
The next step is image segmentation using K-Mean Clustering. This step is done to be able to partition the information into several cluster areas. The results of image segmentation with K-Mean Clustering can be seen in Figure 3 below.

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 6, Number 2, April 2022
DOI : https://doi.org/10.33395/sinkron.v7i2.11333
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Figure 4. Image Segmentation Results with K-Means

The next stage is feature extraction which is useful for extracting information from the identified object characteristics. Feature extraction uses shape and texture feature extraction. The extracted features are then used as boundaries or input values to distinguish objects from one another at the identification stage. In this study, the extraction of shape and texture features was used. For extracting shape features, we use boundary metrics and erraticism. Metric is obtained from the ratio between the area and the circumference of an object. While erraticism is obtained from the comparison between the focal distance of the minor ellipse with the focus of the city chairman ellipse of an object. The following are the results of the feature and texture extraction values implemented in Matlab.



	Ciri	Nilai
1	Metric	0.86776
2	Eccentricity	0.75061
3	Contrast	0.27594
4	Correlation	0.95622
5	Energy	0.38995
6	Homogeneity	0.94627

Figure 5. Results of Value Extraction and Texture

Next is the identification stage with Euclidean distance, an image matching technique that can identify other images that are similar or have similarities. Euclidean distance represents the level of similarity of two images that take into account the distance value from Euclidean, if the smaller the Euclidean distance, the more similar the image. Compositions are carried out by comparing the proximity of the distance values of two variables, namely between the test image and the reference image to find the closest distance value. Figure 6. shows the appearance of the application using Matlab.

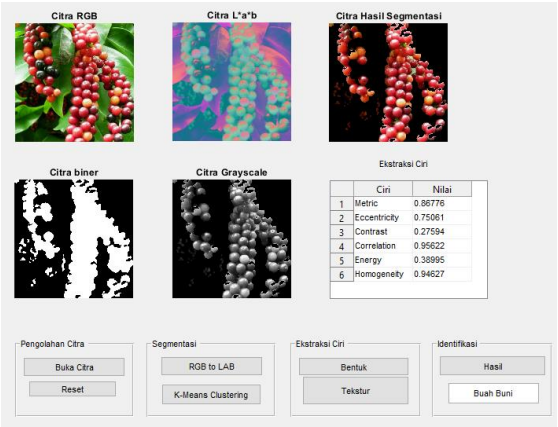


Figure 6. GUI System Identification

Furthermore, based on the results of the disarray network, precision, review and accuracy are calculated using equations (8), (9 and 10). The results of precision, review and accuracy calculations can be seen in Table I.

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 6, Number 2, April 2022
DOI : https://doi.org/10.33395/sinkron.v7i2.11333
e-ISSN : 2541-2019
p-ISSN : 2541-044X

	Ressult
Precision	0.85 %
Recall	0.89 %
Accuracy	0.87 %

Table 1. Results of Precision, Recall and Accuracy Tests

The precision value in Table 1 is 0.85% or 85%. This means that the level of accuracy between the information provided is 85%. While the review got a value of 0.89 or 89%, meaning that the success rate of the system in retrieving information is 89%. Furthermore, for the accuracy of getting a value of 0.87 or 87%, it means that the level of closeness between the predicted value and the actual value is 85%. The percentage classification accuracy criteria are classified into: Good, with a score of 76%-100 percent; Enough, with a value of 56%-75%; Less Good, with a value of 40%-55%, while Less Good, has a value of less than 40%. When viewed from the average accuracy obtained from the identification accuracy percentage of 87%, it is included in the good category. Extraction of shape and texture features can provide information ideally so that it can support the identification process. However, when viewed from the test results, the average accuracy occurs because it is influenced by several factors, including: (1) the shape and texture of the medicinal plant type of buni which is almost the same, making it difficult to distinguish some systems test information; (2) the number of datasets used as training information and test information is still relatively small, so it is not ideal in the learning system; (3) for images with a religious object background can affect the identification results.

DISCUSSIONS

This study examined the buni fruit using the Euclidean distance method with shape and texture extraction. Shape feature extraction uses boundary and unpredictable metrics. This limit is considered to be able to recognize objects based on shape and can distinguish them from other objects. For identification exposition, euclidean distance is used to represent the ranking of two images. This study identifies buni fruit using the Euclidean distance method with shape and texture feature extraction. Extraction of shape features using boundary metric and unpredictability. This boundary is considered to be able to recognize objects based on their shape and can distinguish them from other objects. For identification expositions, euclidean distance is used to represent the level of similarity of two images that take into account the distance value from the euclidean. From the results of the evaluation using a disarray framework by calculating precision, review and accuracy, the value of precision is 85%, recall is 89% and accuracy is 87%. These results indicate that the Euclidean distance and shape and texture feature extraction can identify Buni fruit image objects well. To improve further research several suggestions that can be made. Improvements made are by increasing the amount of training data and test data and being able to use profound learning algorithms to get better feature extraction and identification.

REFERENCES

C. N. Prabiantissa, A. R. Tri, and R. A. Asmara, “Identification System for Natural Batik and Synthetic Batik Based on Image Color Characteristics Using K-Means Clustering Method,” J. Inform. Polynema, vol. 3, no. 2, p. 26, 2017, doi:10.33795/jip.v3i2.10.

D. I. Mulyana, "Optimization of Image Classification Using the Convolutional Neural Network (CNN) Algorithm for Cirebon Batik Image Indonesian," no. 12, pp. 39–46, 2021.

D. Nurnaningsih, D. Alamsyah, A. Herdiansah, and A. A. J. Sinlae, "Identification of Rhizome Types of Medicinal Plants Image by Euclidean Distance Based on Shape and Texture Characteristics," Build. Informatics, Technol. Sci., vol. 3, no. 3, pp. 171–178, 2021, doi:10.47065/bits.v3i3.1019.

H. Prabowo, "Detection of Ripeness Conditions of Citrus Fruits Based on Color Similarities in Android-Based RGB Color Space," J. Elektron. Sis. inf. and Computing., vol. 3, no. 2, pp. 9–19, 2017.

M. Qomaruddin, D. Riana, and A. Anton, “K-Means Segmentation of Tin Leaf Image with Gray Level Co Occurance Matrix Characteristic Classification,” J. Sist. and Technol. Inf., vol. 9, no. 2, p. 223, 2021, doi:10.26418/justin.v9i2.44139.

*name of corresponding author

e-ISSN : 2541-2019
p-ISSN : 2541-044X

N. Sivi Anisa and T. Herdian Andika, "Segmentation-Based Leaf Image Identification System Using K-Means Clustering Method," Aisyah J. Informatics Electr. Eng., vol. 2, no. 1, pp. 9–17, 2020, doi:10.30604/jti.v2i1.22.

R. Enggar Pawening, W. Ja, and far Shudiq, "CLASSIFICATION OF THE QUALITY OF LOCAL ORANGE BASED ON TEXTURE AND SHAPE USING k-NEAREST NEIGHBOR (k-NN) METHOD," Ejournal.Unuja.Ac.Id, vol. 1, no. 1, pp. 10–17, 2020.

S. Hadianti and D. Riana, "Segmentation of Bemisia Tabaci Image Using the K-Means Method," Semin. Nas. Inov. and Trends, p. 2018, 2018.

S. R. Raysyah, Veri Arinal, and Dadang Iskandar Mulyana, "Classification of Coffee Fruit Maturity Levels Based on Color Detection Using Knn and Pca Methods," JSiI (Jurnal System Information), vol. 8, no. 2, pp. 88–95, 2021, doi:10.30656/jsii.v8i2.3638.

S. Sukemi, "Sinta Sukemi," 2021.

SN Saragih, M. Safii, and D. Suhendro, "Implementation of the K-Means Method on Livestock Meat Production," Jurassic (Jurnal Ris. Sist. Inf. and Tek. Inform., vol. 6, no. 1, p. 235, 2021, doi:10.30645/jurasik.v6i1.288.

Z. D. Lestari, N. Nafi'iyah, and P. H. Susilo, "Banana Type Classification System Based on HSV Color Characteristics Using the K-NN Method," Semin. Nas. technol. inf. and Commun., pp. 11–15, 2019.

*name of corresponding author