COMPARISON OF APPLE IMAGE SEGMENTATION USING BINARY CONVERSION AND K-MEANS CLUSTERING METHODS

Siti Nurdiani [1*]; Muhammad Rezki[2]; Rizka Dahlia[3]; Muhammad Ifan Rifani Ihsan[4]; Frieyadie[5]; Siti Fauziah[6]

Program Studi Teknik Informatika, Sistem Informasi
Sekolah Tinggi Manajemen Informatika dan Komputer Nusa Mandiri
www.nusamandiri.ac.id
siti.sxd@nusamandiri.ac.id[1], muhammad.mdk@nusamandiri.ac.id[2], ifan.mii@nusamandiri.ac.id[3], frieyadie@nusamandiri.com[5], siti.suz@nusamandiri.ac.id[6]

Sistem Informasi
Universitas Bina Sarana Informatika
www.bsi.ac.id
rizka.rzl@bsi.ac.id[4]

(*) Corresponding Author

Abstract— Apples are quite popular consumption among the community and have different kinds of shapes and colors. Apples themselves have many nutrients and various vitamins including fat, as well as energy, carbohydrates, protein, vitamin C, vitamin A, vitamin B2, vitamin B1, and many more. Because of the variety of types of apples, it is difficult for people to distinguish between these types of apples. However, with the development of technology and sophistication, it is now possible to classify the types of apples using digital images. This study aims to segment the image of apples by comparing 2 methods at once to find out which method is the best. This process is an initial stage that must be done before classifying. From the comparison results of apple image segmentation with binary conversion methods and k-means clustering, it can be concluded that the best method is k-means clustering. Because it can segment the image of apples almost perfectly.

Keywords: Apples, Image, Binary Conversion, K-Means Clustering

Abstrak—Buah apel cukup populer dikonsumsi dikalangan masyarakat serta memiliki berbagai jenis bentuk dan warna. Buah apel sendiri memiliki banyak nutrisi dan berbagai macam vitamin diantaranya lemak, serta, energi, karbohidrat, protein vitamin C, vitamin A, vitamin B2, vitamin B1 dan masih banyak lagi. Karena beragamnya jenis apel ini membuat masyarakat sulit membedakan jenis-jenis apel tersebut.Namun dengan berkembangnya teknologi dan kecanggihannya saat ini dapat mengklasifikasi jenis-jenis apel dengan menggunakan citra digital. Penelitian ini bertujuan untuk mensegmentasi citra buah apel dengan membandingkan 2 metode sekaligus untuk mengetahui mana metode yang paling baik. Proses ini merupakan tahap awal yang harus dilakukan sebelum melakukan klasifikasi. Dari hasil perbandingan segmentasi citra buah apel dengan metode konversi biner dan *k-means clustering* dapat disimpulkan bahwa metode yang paling baik adalah *k-means clustering*. Karena dapat mensegmentasi citra buah apel dengan hampir sempurna.

Kata Kunci: Buah apel, Citra, Konversi Biner, K-Means Clustering

INTRODUCTION

Apples are a type of fruit that has a variety of colors and varieties. Various types can also be known from the color of the apple itself. The fruit, which has the Latin name Malus Domestica, has good levels of vitamins for the body. Some types of apples that we often encounter include red apples, crimson apples, fuji apples, poor apples, and many more. Sweets, apple chips and drinks are processed products from apples besides being able to be consumed directly. Of course, if you want to process apples, you need knowledge about the ripeness of apples that are suitable for processing so that the taste and quality that will be produced are maintained properly (Ciputra et al., 2018).
Apples themselves have many nutrients and various vitamins including fat, as well as energy, carbohydrates, protein, vitamin C, vitamin A, vitamin B2, vitamin B1, and many more (Ciputra et al., 2018). Apples also contain carotene, carotenoids

have activity as vitamin A and antioxidants that are beneficial to counteract free radical attacks that m Causes of right all degenerative diseases (Qisti et al., 2020).

Apples become one fruit that is favored by the people of Indonesia. The average consumption of apples in Indonesia to 1,1 kg per capita per year according to the central agency statistics in 2006. The compound of polyphenols contained in apples can be self-cleansing that prevent plaque formation through biochemical reactions by catechins (Prianggara et al., 2020).

The various types of apples make it difficult for people to distinguish between the types of apples themselves. However, with the development of technology and sophistication, it is now possible to classify the types of apples using digital images. Image is another term for images as one of the multimedia components that play a very important role as a form of visual information (Premana et al., 2020). Digital images are arrays or *arrays* that contain original or complex values represented by a row of bits (Zainuddin et al., 2017). Image acquisition, image pre-processing (segmentation), feature or feature extraction, training, and testing are the stages in the image classification process (Safitri, Nurdiani, Riana, & Hadianti, 2019). Image processing (image processing) is a system that was initially carried out to improve image quality, but as the world of computing develops, image processing cannot be separated from the field of computer vision (Nafi'iyah & Fatichah, 2017).

In a study (Safitri et al., 2019) on the classification of apple species using the Order 1 and Multi-SVM methods but with public datasets taken from the *Kaggle* website. While in this study the dataset used is a personal dataset that is taken directly by the researcher and uses a comparison with 2 segmentation methods.

This study aims to segment the image of apples by comparing 2 methods at once to find out which method is the best. This process is an initial stage that must be done before classifying.

MATERIALS AND METHODS

The design of apple fruit image processing can expressed in Figure 1:

1. Image Capture

In this study, the dataset used is private and not a public dataset. The image dataset of apples photographed using a cellphone camera at the same distance and time. The amount of data used is 20 images of apples.

2. Original Image converted to Grayscale

Following the image acquisition process, then the image processed by means converted from RGB to grayscale. A grayscale image is a digital image that has only one channel per pixel, in other words, the value of the Red = Green = Blue part. This value used to indicate the level of intensity (Firmanto, Rikasanti, Bramanto, & Putra, 2019). The intensity level is shown starting from black, gray to white (Raihana Salsabila Darma, Adiwijaya, Andrian B Suksmono, 2021) . With the following equation:

$$Gray = 0.2125 \times R + 0.7154 \times G + 0.0721 \times B$$

Figure 1. Binary conversion segmentation image processing

3. Binary Conversion Segmentation

Segmentation is the process of dividing the image into multiple parts (Sinaga, 2017). Image segmentation is the processof getting information from the image and understanding the characteristics of the image completely used for the process to the next level, for example, the object identification process (Alvini & Dewi, 2021)). After grayscale, the image is segmented using binary conversion. The binary conversion is an image that only has two intensity values, namely 0 (black) and 1 (white).

In addition to using the binary conversion segmentation method, images also segmented using the K-Means Clustering method as described in Figure 2:

1. Original image converted to LAB

In this process, the original image of the apple in RGB form is converted or converted into the LAB

color space. R is the color L * a * b *, L * represents the brightness, and a * and b * are the chromaticity coordinates. a * and b* are the color directions: +a* is the red axis, -a' is the green axis, +b* is the yellow axis, and -b* is the blue axis. This color space describes all colors visible to the human eye, where L indicates light, A indicates the position from red (red) to green (green) and B indicates the position between yellow (yellow) and blue (blue) (Alamsyah & Pratama, 2019).

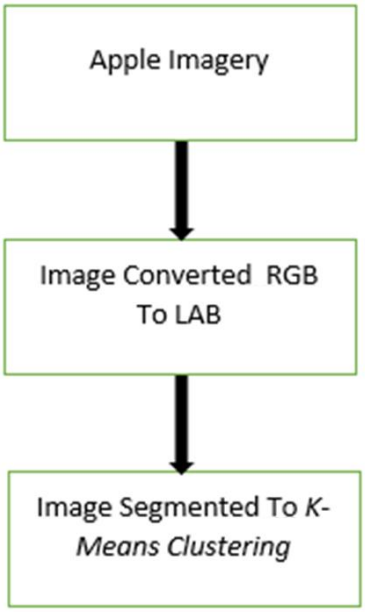

Figure 2. Image processing of K-Means Clustering segmentation

2. *K-Means Clustering Segmentation*

Clustering is a data mining technique that is used to get groups of objects that have the same characteristics in large enough data (Suriani, 2020). K-Means Clustering is a non-hierarchical data clustering method that divides existing data into several groups based on the same characteristics (Rahmah, 2020). K-Means is a fairly simple clustering algorithm that partitions data into several clusters, k-means is quite easy to implement and has also been widely used (Sibuea & Sapta, 2017). It based on the ease of implementation and speed of the process. The following is the equation for calculating the *K-Means Clustering*:

$$d\ (x,y) = \sqrt{(x1 - x1)^2)} + (y1 - y2)\quad ^2 \ldots (yn - yn)^2$$

RESULTS AND DISCUSSION

1. Image Capture Process

The image is taken using a *cellphone* camera. The distance and time of taking the image of apples are the same for each fruit. The image below is 10 images of apples that have been *captured.*

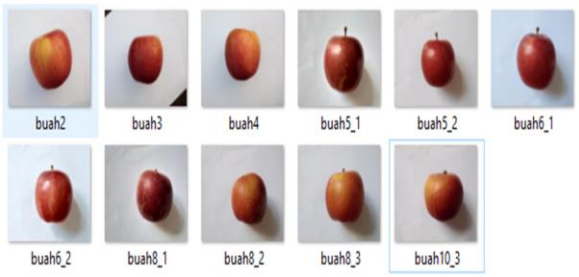

Figure 3. Apple Fruit Dataset

2. Original Image converted to *Grayscale*

In the image below is an image of an apple that has been converted into a gray or grayscale color space. The channel value red = green = blue has one channel value in each pixel of the gray level image. The intensity level is shown starting from black, gray to white (Raihana Salsabila Darma, Adiwijaya, Andrian B Suksmono, 2021). With the following equation:

Gray = 0.2125 x R + 0.7154 x G + 0.0721 x B

The image will change color from the original to gray as shown below.

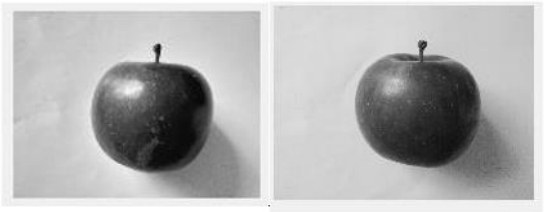

Figure 4. The original image is converted to grayscale

3. Original Image converted to Grayscale

In the image below, you can see the image of an apple after being converted to grayscale and then segmented into binary conversion. Converting from a grayscale image to a binary image aims to obtain the edges of the object, an image with a binary image type, each pixel is only expressed by a value of two possibilities (namely values 0 and 1) so that it can be analyzed easily, as for how to apply the value known as the threshold value(Firmanto et al., 2019). The threshold is a segmentation technique commonly used for images with an important difference in intensity values between the background and the object (Satun & Pandiangan, 2020). Value The resulting image will look black and white. The shape of the image of an apple can be seen from the white pattern.

Figure 5. Image segmentation using binary conversion

4. Original image converted to LAB

In Figure 6 below, it can be seen the results of changing the image of the apples which changed the color space from RGB to LAB, the apple shape area is dark green compared to other areas.

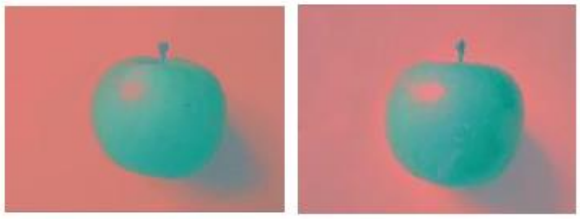

Figure 6. Original image converted to LAB

5. K-Means Clustering Segmentation

Figure 7 below can see results of the image changes apples segmented using the *k-means clustering* clearer shape and its area. The shape of the red apple very clearly separated from the background. With the following calculation equation:

$$d\ (x,y) = \sqrt{(x1-x1)^2) + (y1-y2)\quad ^2 \ldots (yn-yn)^2}$$

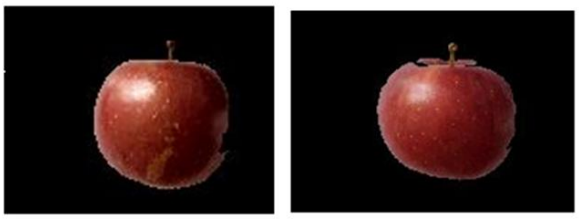

Figure 7. Image segmentation using *K-Means Clustering*

Figure 8 below is the result of the comparison of apple image segmentation using the binary conversion method and the k-means clustering method. It can see that the difference is quite significant where the image of apples segmented using binary conversion area formed imperfectly and there are still areas that have not been segmented properly. While the image of apples segmented with K-Means Clustering the results are good with the condition of the area of the image is perfectly formed separated between the background and the image.

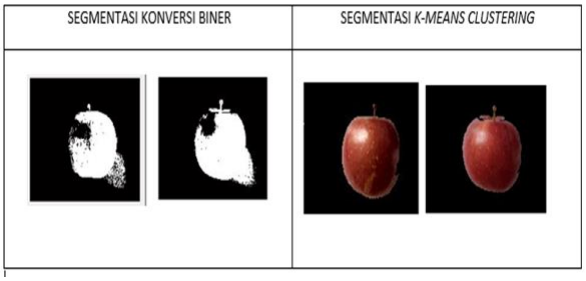

Figure 8. Comparison of apple image segmentation

In the apple fruit image segmentation trial using the binary conversion method with 10 apple images, there were only 5 well-segmented images or about 50%, while 50% failed to segment properly.
Then in the trial of apple image segmentation using the *k-means clustering* method with 10 apple images, 8 images were well-segmented or about 80%, and the remaining 2 apple images or about 20% could not be segmented properly.

CONCLUSION

In this study, it can conclude that segmentation using the *k-means clustering* method has better results than using the binary conversion method. The results of the image of apples segmented using the *k-means clustering* method are more perfectly round. While the segmented with binary conversion round is not perfect and there are still many empty parts that are not covered. With the details of the total trial of apple image segmentation using the binary conversion method with 10 apple images, there are only 5 images that are well segmented or about 50%, while 50% fail to be well segmented. Then in the trial of apple image segmentation using the *k-means clustering* method with 10 apple images, 8 images were well-segmented or about 80%, and the remaining 2 apple images or about 20% could not segmented properly.

REFERENCE

Alamsyah, D., & Pratama, D. (2019). Segmentasi Warna Citra Bunga Daisy dengan Algoritma K-Means pada Ruang Warna Lab. *Jurnal Buana Informatika*, *10*(2), 153. https://doi.org/10.24002/jbi.v10i2.2458

Alvini, S., & Dewi, M. P. (2021). Penerapan Pohon Rentang Minimum pada Graph dalam Segmentasi Citra. *UNP Journal of Mathematics*, *4*(2), 57–61.

Ciputra, A., Susanto, A., & dkk. (2018). Dengan Algoritma Naive Bayes Dan Ekstraksi Fitur Citra Digital. *Simetris : Jurnal Teknik Mesin, Elektro Dan Ilmu Komputer*, *9*(1), 465–472.

Firmanto, B., Rikasanti, E., & Putra, A. B. W. (2019). Optimasi Hasil Akuisisi Obyek Wajah

Menggunakan Metode Binary Large Objects. *Semnas SENASTEK Unikama 2019*, *2*, 826–840. https://conference.unikama.ac.id/artikel/index.php/senastek/article/view/312

Nafi'iyah, N., & Fatichah, C. (2017). Fuzzy self organizing map untuk proses thresholding pada citra dental panaromic. *Seminar Nasional Sistem Informasi*, *September*, 511–524.

Premana, A., Bhakti, R. M. H., & Prayogi, D. (2020). Segmentasi K-Means Clustering Pada Citra Menggunakan Ekstrasi Fitur Warna dan Tekstur. *Jurnal Ilmiah Intech : Information Technology Journal of UMUS*, *2*(01). https://doi.org/10.46772/intech.v2i01.190

Prianggara, ferdian wahyu, Setiawan, ahmad bagus, & Farida, intan nur. (2020). *Identifikasi Jenis Buah Apel Berdasarkan Ekstraksi Bentuk dan Warna*.

Qisti, N., Nurwidah, A., Padapi, A., & Haryono, I. (2020). Analisa Kelayakan Usaha Pembuatan Selai Apel di UMS Rappang Store. *MALLOMO: Journal of Community Service*, *1*, 22–29.

Rahmah, S. A. (2020). Klasterisasi Pola Penjualan Pestisida Menggunakan Metode K-Means Clustering (Studi Kasus Di Toko Juanda Tani Kecamatan Hutabayu Raja). *Djtechno : Journal of Information Technology Research*, *1*(1), 1–5.

Satun, H., & Pandiangan, M. (2020). Segmentasi Citra Untuk Pencarian Kode Warna Cat Menggunakan Metode Thershold Hsv. *Bulletin of Information Technology (BIT)*, *1*(3), 134–143.

Sibuea, F. L., & Sapta, A. (2017). Pemetaan Siswa Berprestasi Menggunakan Metode K-Means Clustering. *JURTEKSI (Jurnal Teknologi Dan Sistem Informasi)*, *4*(1), 85–92.

Sinaga, A. S. R. (2017). Implementasi Teknik Threshoding Pada Segmentasi Citra Digital. *Jurnal Manajemen Dan Informatika Pelita Nusantara*, *1*(2), 48–51.

Suriani, L. (2020). Pengelompokan Data Kriminal Pada Poldasu Menentukan Pola Daerah Rawan Tindak Kriminal Menggunakan Data Mining Algoritma K-Means Clustering. *Jurnal Sistem Komputer Dan Informatika (JSON*, *1*, 151–157. https://doi.org/10.30865/json.v1i2.1955

Zainuddin, M., Sianturi, L. T., & Hondro, R. K. (2017). Implementasi Metode Robinson Operator 3 Level Untuk Mendeteksi Tepi Pada Citra Digital. *Jurnal Riset Komputer (JURIKOM)*, *4*(4), 1–5.