Content based Image Retrieval To Identify Medicinal Plants using Shape and Color Features

Jumi[1], Achmad Zaenuddin[2],Tedjo Mulyono[3*)]
Departement of Business Administration[1,2]
Departement of Civil Engineering[3]
Politeknik Negeri Semarang,
Semarang, Indonesia[1,2,3]

Nur Hayati[4]
Departement of Information Technology[,4]
Institut Teknology and Sains Nahdlatul Ulama
Pekalongan, Pekalongan, Indonesia[4]

***Abstract* - This is due to the properties possessed by these plants. There are thousands of species of medicinal plants with efficacy as efficacious herbal medicines with low side effects. Medicinal plants have different physical characteristics. Species of medicinal plants can be distinguished from the shape and color of the leaves. The main feature that becomes the dominant feature to be able to distinguish each species of medicinal plants is the feature of color and leaf shape. These features become the main key in the identification of medicinal plants.**

Identification is done by comparing the similarity of images of medicinal plant leaves using color and shape feature values. The similarity can be determined through differences in the content of medicinal plant leaf images, namely the value of shape and color features between the query image and the database image. The closer to zero the difference, the higher the level of similarity. The level of similarity will affect the accuracy of image recognition at the time of identification. In this study, an analysis of the accuracy of the identification of medicinal plant leaves was carried out and the computational time measurement for its identification was carried out. Improved identification accuracy by weighting the value of shape and color features. Extraction of medicinal plant leaf images uses invariant moment for shape features and color moment for color features. Preprocessing before extraction using Grayscale, resize, Histogram Equalization and Edge Enhancement. Rice image data clustering using K-Means clustering. The results showed that the identification accuracy with test data of 500 images of medicinal plant leaves reached more than 80% on a weighting scheme of Ws (weighted Shape) = 60% and Wc (weighted color) = 40% with an average computation time of less than 5 milli -second on 15 clusters.

Keywords : Weighted , Feature, Retrieval, Moment, K-Means

1. INTRODUCTION

Indonesia is a country that is rich in plants with various types and different benefits. Each plant species has a leaf shape and color that varies. Indonesia has a wealth of medicinal plants that have the potential to be developed. Indonesia has more than 35,000 plant species with more than 2000 species being types of herbal medicinal plants [1]. The large number of herbal medicinal plants and the lack of public knowledge about types and herbal medicinal plants make it difficult for people to distinguish the species and efficacy of these herbal medicinal plants, so many people prefer to use chemical medicine. To provide information to the public, a system for the introduction of herbal medicinal plants is needed that is able to identify and recognize herbal medicinal plants.

Identification can be affected by the accuracy of various variables and the accuracy in the process of calculating feature values. [5]

The information obtained can be in the form of digital images which are then analyzed and processed by the system. The system identifies leaf images from herbal medicinal plants and recognizes a pattern or characteristic of the object. Research on identification in an image has long been developed, one of them is by distinguishing the texture of the image. Other studies have shown that regardless of feature weights and clustering, it can produce a similarity accuracy value of more than 75% [7]. Retrieval by weighting the texture feature using DCT (DiscreteCosine Transform) has also been carried out in CBIR-based image retrieval [16]. The features' values will be used as the coefficient of the image's value to be used as the basis for the retrieval.

Image texture can be distinguished by density, regularity, uniformity, and roughness [2]. Because computers cannot distinguish textures like human vision, texture analysis is used to determine the pattern of a digital image. Texture analysis will produce a value of texture characteristics or characteristics which can then be processed by a computer for the classification process [3]. Medicinal plants are one type of plant that has a very important role in human life, in addition to the stems, flowers, leaves, fruits, and roots that are most often used in the identification process of medicinal plants, there are leaves. However, physical characteristics in the form of color are considered not so significant in determining the type of leaf. This is because almost all types of leaves have a dominant green color. Meanwhile, to obtain shape features, sometimes it is difficult to collect leaf data as a whole, especially for leaves that have a large scale, so that the combination of shape, color and leaf texture features is a more appropriate feature to use in the identification of leaves of herbal medicinal plants.

The leaves that will be studied in this study are the types of leaves of traditional medicinal plants which include herbal plants. Image identification in other studies was investigated by comparing the feature values between the query image and the image in the database [4]. In this study, researchers want to create a system that can identify traditional medicinal plants that are around to help ordinary people in the identification process. The identification process is carried out by recognizing the shape, color and texture of the leaf image that is applied to a type of leaf of a traditional medicinal plant and matching it with data from the database of medicinal plant leaf images. The recognition process is carried out by extracting features using the Invariant moment, Color Moment and Statistical Texture methods. The features of each leaf are used for the identification process, there are the Euclidean Distance and K-Means Clustering methods.

The quality of a retrieval system depends on the features used to describe image content [6]. By using a computer vision-based approach that was carried out on the classification of six types of rice with a total of 21 features of rice grains based on seven characteristics on color and fourteen morphological characteristics and using the Neural network method, the accuracy reached up to 88.30% [8]. It is different with the detection of bananas in the natural environment using color and texture features in a study using a regular red-green-blue color camera. 120 photos under different illumination conditions were selected as the test set and obtained a single-scale average detection rate of 89.63%. This shows that the method used can detect bananas in plantations under different lighting and occlusion conditions [9].

Leaf classification based on texture characterization, shape, and color properties using the Fuzzy Relevance Vector Machine (FRVM) for leaf type characterization was also carried out so that it could accurately predict leaf types from the input leaves given the images. The experimental analysis showed better results such as 99.87% accuracy. This value indicates an increase in the value of the reference study [10]. However, identification of oil palm leaves using 41 features with principal component analysis (PCA) and Artificial Neural Network (ANN) methods resulted in an accuracy of 99.67% which indicates a decrease in accuracy from the previous case. Therefore, it is necessary to conduct an experiment to justify the appropriate method for the dataset used in this study [12].

Plants are believed to be able to play a very important role in maintaining the ecological balance of the Earth by providing us with an atmosphere, food, shelter, medicine, fuel and a source of oxygen. [14]. This study aims to identify types of medicinal plant leaves based on shape, color and texture in determining the types and efficacy of medicinal plants. Another objective is to analyze the identification of leaf species based on an image using a feature value weighting scheme and calculating the computation time of tracing after clustering using the K-Means technique. The K-Means algorithm is used to determine the position of the cluster in each image by first calculating the distance between the image and all the centroids using the Euclidean distance method.

2. RELATED WORKS

2.1 Color Moments

Color is the result of the perception of light in the spectrum of the visible region of the retina of the eye, and has a wavelength between 400nm to 700nm. Visualization of RGB color money is a cube like in Fig. 1, which has 3 axes to represent the color R (red) or red, G (green) or Green and B (blue) or blue. One of the opposite corners of the base represents black when the value R=G=B=0, while the opposite top corner represents white when the value R=G=B=255 (8 bit color system).

Fig. 1. RGB color components as color intensity vector

The RGB color space does not allow the separation of luminance information from color. However, in the HSV color space, one can separate them. HSV represents three parts of the color space, namely hue representing color, saturation of color purity and pixel brightness value [14]. Colors are presented on the basis of three basic colors, where each color can be reproduced by mixing a set that corresponds to the three basic colors. The number of colors depends on the three color component vectors in the three-dimensional space of the coordinate system. The set of all colors forming a vector space is called a color space or color model. Color can be described as a representation of the color metric space so that a color difference can be calculated using the difference in distance through the Euclidean distance. The transformation equations shown by equations 1 to 4 are used for the conversion of vector values from RGB to HSV.

$$r = \frac{R}{(R+G+B)}, g = \frac{G}{(R+G+B)}, b = \frac{B}{(R+G+B)} \quad (1)$$

$$V = \max(r, g, b) \quad (2)$$

$$S = \begin{cases} 0, & if\ V = 0 \\ 1 - \frac{\min(r,g,b)}{v}, & V > 0 \end{cases} \quad (3)$$

$$H = \begin{cases} 0, if\ S = 0 \\ \frac{60*(g-b)}{S*V}, if\ V = r \\ 60 * \left[2 + \frac{b-r}{S*V}\right], if\ V = g \\ 60 * \left[4 + \frac{r-g}{S*V}\right], if\ V = b \end{cases} \quad (4)$$

$$H = H + 360\ if\ H < 0$$

In equations 2, 3 and 4 are the equations for converting from RGB to HSV. The HSV (Hue, Saturation, Value) color model is depicted in Fig. 2.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 10 Issue 12, December-2021

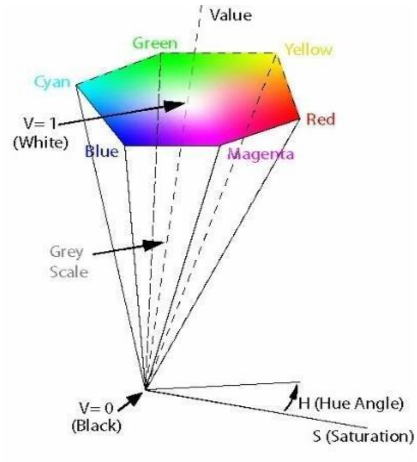


Fig. 2. HSV Hexcone color model

Color Moments is a method used to distinguish images based on their color features. The theoretical basis of this method is to assume that the color distribution in an image can be expressed as a probability distribution. Color Moment is a fairly good method of recognizing color characteristics. This method uses three main moments from the color distribution of the image, namely the mean, standard deviation, and skewness, so that this method produces three values for each color component [2]. Each color component, namely HSV (Hue, Saturation and Value) has 3 moments. The calculation of the three moments uses equation 5, equation 6 and equation 7

Preprocessing with histogram equalization in Fig. 4 and Fig. 5 produces a more contrasting image. This image is the input image for color feature extraction. The extraction method used is color moment. This method is able to distinguish images based on their color features [15]. Color Moment is a pretty good method of recognizing color characteristics. This method uses three main moments of the image color distribution, there are the mean, standard deviation, and skewness [13]. Each color component, there are HSV (Hue, Saturation and Value) has 3 moments. The three moments are calculated using equation (5), equation (6) and equation (7).

a. Moment 1 – Mean :

$$Ei = \sum_{N=-1}^{j=1} \frac{1}{N} P_{ij} \quad (5)$$

Mean : can be said as the average color value in the image.

b. Moment 2 – Varians :

$$\sigma_i = \sqrt{\frac{1}{N}\sum_{N}^{j=1} (P_{ij} - E_i)^2} \quad (6)$$

$$\sigma i = \sqrt{(\frac{1}{N}\sum_{j-1}^{N} (Pij - Ei)^2)}$$

Standard deviation: the range of spread of the data from the mean

c. *Moment 3 – Skewness* :

Skewness is a measure of the asymmetry of the data around the mean, calculated by equation (7). Skewness is a measure of the degree of asymmetry of color distribution.

$$S_i = \sqrt[3]{\frac{1}{N}\sum_{N}^{j=1} (P_{ij} - E_i)^3} \quad (7)$$

The distance from the color distribution of the query image to the database image is calculated by equation (8).

$$D_{color}(H,I) = \sum_{i=1}^{r} W_{i1} |E_i^1 - E_i^2| + W_{i2} |\sigma_i^1 - \sigma_i^2| + W_{i3} |S_i^1 - S_i^2| \quad (8)$$

E	=	The average color value of the image (Mean).
(H,I)	=	Two images compared
I	=	HSV Color Component Index (H = 1, S = 2, V = 3)
R	=	Number of Indexes (3)
σ	=	Root of the variance (Standard Deviation).
S	=	A measure of the degree of asymmetry (Skewness).
N	=	The total number of pixels in the image.
J	=	Pixel order.
Wi	=	The weight of each moment.
Pij	=	The value of the *i* color component at the *j* pixel

2.2 *Invariant Moment*

Shape feature extraction has been conducted by using the moment invariant method. This method is used because it is not susceptible to image changes caused by Rotation, Scale and Translation (RST) [13]. Invariant Moment will produce seven invariant moment values that are constant with respect to RST [3]. In the moment invariant analysis, it is also possible to analyze the form of the moment set of a function *f(x,y)* of two variables which are defined as follows [2]:

The moment order *(p+q)* for a continuous function *f(x,y)* with p,q = 0, 1, 2, … is defined as follows:

$$m_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^p x^q f(x,y)dxdy \quad (9)$$

The formula for the central moment is written as follows:

$$\mu_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \underline{x})^p (y - \underline{y})^q f(x,y)dxdy \quad (10)$$

With

$$\underline{x} = \frac{m_{10}}{m_{00}} \; dan \; \underline{y} = \frac{m_{01}}{m_{00}} \quad (11)$$

And for digital images

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 10 Issue 12, December-2021

$$\mu_{pq} = \sum_x \sum_y (x - \underline{x})^p (y - \underline{y})^q f(x,y) \quad (12)$$

The region is described by seven invariant moments that can be searched from the central moment: 00, 10, 01, 20, 11, 02, 30, 12, 21, 03 , and by defining normalized central moments, there are $pq = \mu pq$ / 00, where = 0.5 ($p + q$) + 1 for $p + q = 2, 3$, and so on, which finally gets seven invariant moments, the seven invariant moments are shown in equation (13) to equation (19)

$$M_1 = \eta_{20} + \eta_{02} \quad (13)$$

$$M_2 = (\eta_{20} + \eta_{02})^2 + 4\eta_{11}^2 \quad (14)$$

$$M_3 = (\eta_{30} - \eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \quad (15)$$

$$M_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2 \quad (16)$$

$$M_5 = (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} - \eta_{03})^2] \quad (17)$$

$$M_6 = (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}) \quad (18)$$

$$M_7 = (3\eta_{21} - \eta_{30})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + (3\eta_{12} - \eta_{30})(\eta_{21} - \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \quad (19)$$

The values of these seven invariant moments do not change with rotation, translation and scale [13].

2.3 Image Similarity Accuracy

Determination of the accuracy level of similarity of medicinal plants for tracing the identity of medicinal plants is to calculate the distance normalization of the two methods above. The measurement of the accuracy of similarity using the precision formula as follows:

Accuracy = (the amount of data deemed relevant by the user / the amount of data retrieved by the system) x 100%

3. METHOD

This study uses a method by combining color and shape features. The method used consists of three stages, namely data acquisition through object capture, image preprocessing, clustering, matching similarity between the input image and the image stored in the database for recognition and identification as shown in Fig. 3. This study uses a combination method of shape and color feature values, then weighting the two feature values and clustering which was not found in previous studies, namely the recognition method based on color and shape features for the identification process of medicinal plants through leaf imagery.

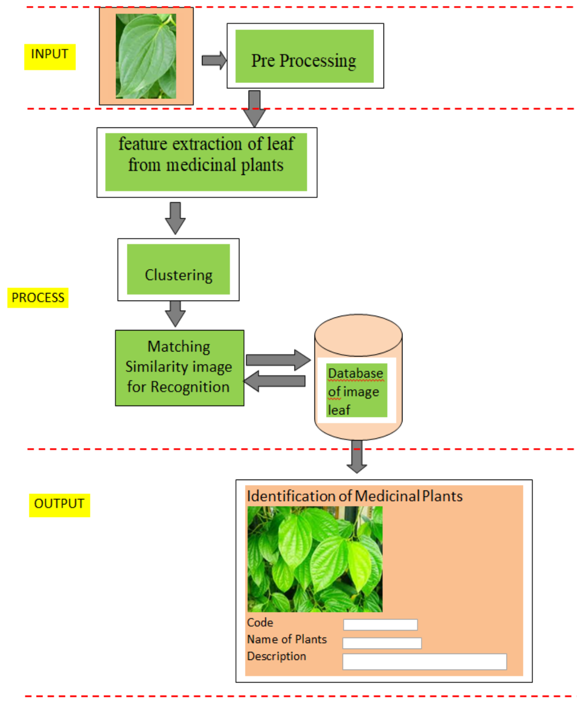


Fig. 3. Overview of Image Identification Stages

The details of the stages of this research are as follows:

3.1 Preprocessing

Preprocessing is the stage of image quality improvement before feature extraction with the aim of increasing the accuracy of the image feature extraction results. There are differences in the preprocessing of color feature extraction with shape feature extraction. The difference in preprocessing is to get a quality image before feature extraction.

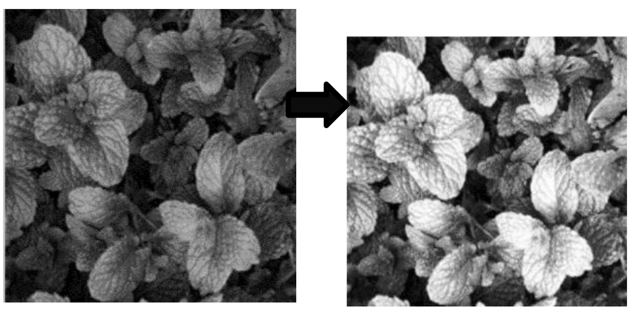
3.1.1 Color Feature Preprocessing

a. *Resize*

The larger the size of the image, the longer the extraction time, so a resize step is needed to speed up the computational process. At this stage, the image is resized to 200 x 200 pixels.

b. *Histogram Equalization*

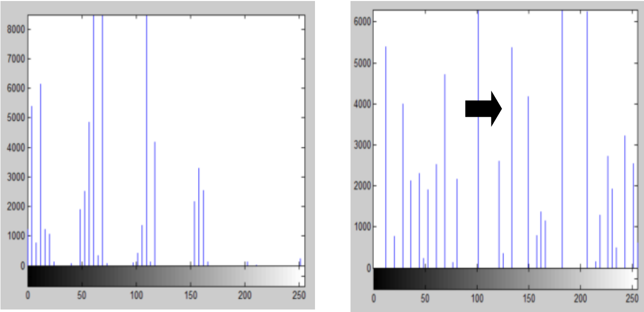
At this stage, the histogram is smoothed so that the image quality becomes more contrasting to get a quality color feature value. The results of histogram smoothing can be seen in Fig. 4 and Fig. 5.



(a) (b)

Fig. 4. Comparison of images before and after the histogram process, (a) Original Image, (b) Histogram Equalization (HE) Result Image

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 10 Issue 12, December-2021



(a) (b)

Fig. 5. Histogram comparison, (a) Original Image Histogram, (b) HE Result image histogram

3.1.2 Preprocessing Shape Features

a. *Resizing*

The size of the image will affect the length of time for feature extraction, the smaller the size, the faster the computation of feature extraction. At this stage, the image is resized to 200 x 200 pixels.

b. *Grayscale*

Especially in calculating the shape features of the color elements, it is not taken into account, so that the image changes to grayscale so that the computation process becomes faster.

c. *Edge Enhancement dan Histogram Equalization (HE)*

The edge enhancement stage will produce a new image with clearer or sharper object edges, so that the shape of the image will be clearer. This stage uses the convolution method with the Sobel operator [3]. The results of edge sharpening can be seen in Fig. 6.

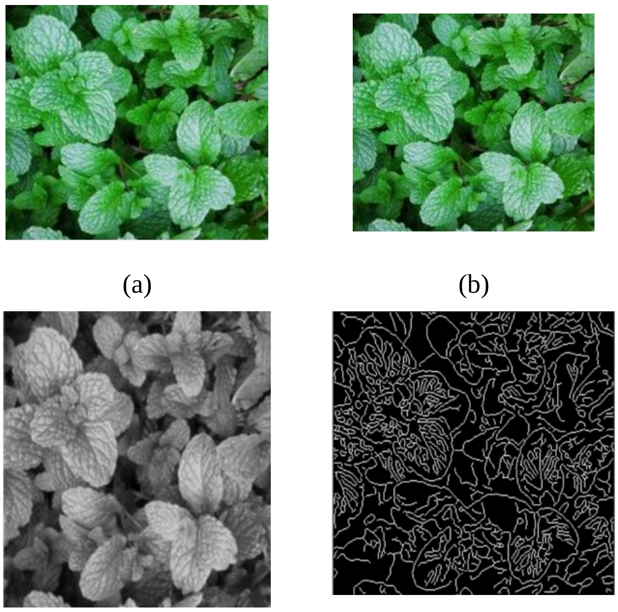


Fig. 6. Preprocessing Stages of Shape Features, (a) Original Image (b) Resize Image,
(c) Grayscale Image, (d) Image after Edge Enhancement Process

3.2 Feature extraction

Color and shape feature extraction is used in the image after going through the preprocessing stage. Color feature extraction uses color moment and shape feature extraction uses invariant moment. The Color moment uses three main moments from the color distribution of the image, namely the mean, standard deviation, and skewness, so this method produces three values for each color component [11]. The Color Moments method is a method used to distinguish images based on their color features [12]. The Color moment assumes that the color distribution in an image can be expressed as a probability distribution. Invariant moment produces feature values that are not susceptible to image changes caused by Rotation, Scale and Translation (RST) [11]. Analysis of the set of moments of a function *f(x,y)* of two variables can also be done [2].

3.3 Feature Weighting

This study uses different weighting feature values for identification accuracy analysis. The varying weighting of feature values will result in different levels of identification accuracy. The weighting is carried out with four variations, namely the *W1* scheme (0.30, 0.70), which is a 30% weighting scheme for color feature weights and 70% shape feature weight. The *W2* scheme (0.6, 0.4) is a weighting scheme of 60% weight feature color and 40% weight feature weight. The *W3* scheme (0.70, 0.30) is 70% weight feature color and 30% weight feature shape. And in the *W4* scheme (0.4, 0.6) there is a 40% weighting scheme for color feature weights and 60% shape feature weight.

3.4 Clustering

Clustering data is useful to speed up the identification process. In this study, the K-Means Clustering and Euclidean distance methods were used as a measurement of image similarity. The similarity measurement equation uses equation (20).

$$D(Q,M) = \sqrt{\sum_{n=1}^{k} (Q_n - M_n)^2} \quad (20)$$

Where Q_n and M_n are features of the query image and database image on the *n* dimension.

Image similarity is determined from the difference in the value of the image features identified with the image in the database. The difference in feature values that is close to zero has the highest level of similarity.

4 RESULT AND DISCUSSION

The test in this study used 10 species of medicinal plants with a total of 500 leaf images measuring 200 x 200 pixels.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 10 Issue 12, December-2021

Table 1. Data Acquisition Of Medicinal Plant

Image	Name of Plants	Image	Name of Plants
	Mint		Binahong
	Sirih Hijau		Sirih Merah
	Sambiloto		Brotowali
	Sambung nyowo		Dadap Srep
	Kelor		Daun Dewa

In Table 1 are some examples of image data and names of medicinal plants used for testing.

4.1 Identification Results

The identification test was carried out with variations in the number of clusters and variations in the percentage of feature weights. The test results with 60% weighting variations for shape feature weights and 40% color feature weights with a variation of 10 clusters are shown in Fig. 7.

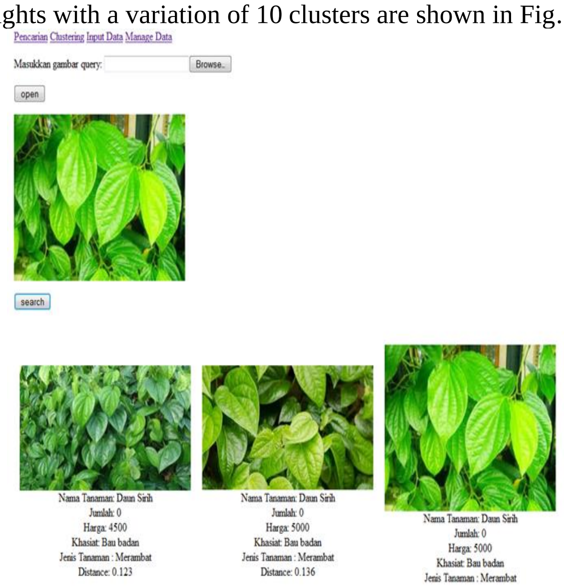


Fig. 7. Results of Tracing the identity of Medicinal Plants using the W2 Scheme

In Fig. 7. the results of the identification of medicinal plants are shown using the similarity value with the Euclidean distance method. Retrieval has a similarity rank of 1 to 10 from the image database of medicinal plant leaves. Tests with variations in weighting and variations in the number of clusters produce different retrieval outputs on the same data.

4.2 Discussion

At this stage, retrieval accuracy is calculated on the image database of medicinal plant leaves. Furthermore, analysis of retrieval accuracy was carried out on the image database of medicinal plants before and after clustering with the number of clusters varying from 3 to 15 clusters.

4.2.1 Before Clustering

Analysis of retrieval accuracy at this stage uses a database of medicinal plant leaf images before clustering with weighting variations on shape and color features. The accuracy value of retrieval results with the same feature weights and varying feature weights can be seen in Table 2.

Table 2. Percentage Of Retrieval Accuracy With Same Feature Weight And Varied Feature Weight In Image Database Of Medicinal Plant Leaves.

Image of Name	Equal Weight	W1	W2	W3	W4
D1	75,55%	78,59%	61,33%	84,38%	86,75%
D2	74,67%	79%	60,67%	83,87%	87%
D3	75,05%	79,45%	60%	83,33%	87,10%
D4	74,69%	79,15%	59,33%	86,67%	87,50%
D5	75,65%	78,75%	60,67%	84,38%	86,67%
D6	75%	78,50%	60%	84%	86,36%
D7	75,76%	79,45%	59,33%	83,87%	86,67%
D8	74,89%	78,65%	60,67%	83,33%	87,33%
D9	74,87%	79,55%	61,33%	84%	86%
D10	75,35%	79,25%	60,67%	84,38%	87,33%

Table 2 shows that the retrieval accuracy with *W4* variation produces the highest average accuracy and *W2* weight variation produces the lowest accuracy.

3.2.2 *After Clustering*

To speed up the process of identifying medicinal plants, image data is grouped. The grouping is done based on the similarity which the results will further narrow the search space for image data information. Init grouping will shorten retrieval time and improve retrieval accuracy. The results of the retrieval accuracy test with variations in weight and number of clusters can be seen in Table 3.

Table 3. Percentage Of Retrieval Accuracy With Variation In Feature Weight And Variation In Number Of Clusters

Number of cluster	Equal Weight	W1	W2	W3	W4
3	73%	76%	63%	82%	85%
4	73%	75%	64%	82%	86%
5	73%	76%	64%	83%	88%
6	74%	76%	64%	84%	88%

7	75%	78%	66%	84%	89%
8	74%	77%	65%	84%	89%
9	75%	77%	64%	84%	89%
10	76%	80%	67%	86%	89%
11	74%	79%	66%	84%	88%
12	74%	77%	66%	85%	89%
13	74%	76%	65%	83%	85%
14	73%	75%	64%	82%	85%
15	73%	75%	64%	82%	85%

Based on Table 3, it is shows that the weighting variation of *W4* with 60% of color features, 40% of color features and variations in the number of clusters 10 resulted in the highest retrieval accuracy of more than 85%. While the weighting variation of *W2* with 40% of the weight of color features and 68% of the weight of color features with variations of the number of clusters 3 produces the lowest accuracy, which is less than 64%.

The graph of retrieval accuracy in the rice image database uses the same and varied feature weights, namely *W1*, *W2*, *W3* and *W4* with variations in the number of clusters ranging from 3 to 15 as shown in Fig. 8.

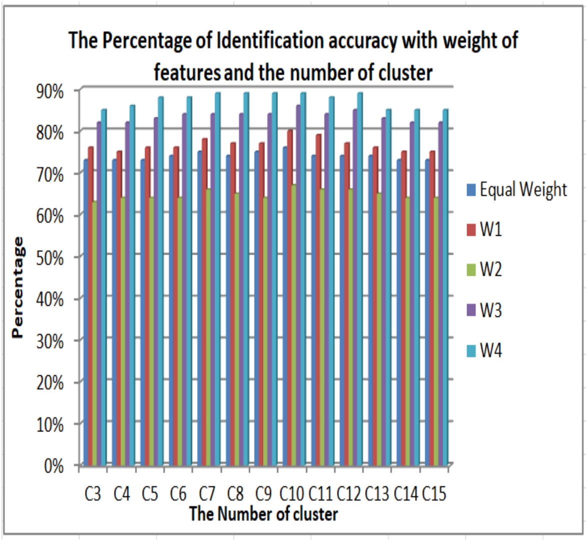


Fig. 8. Graph of Retrieval Accuracy with Variations in Feature Weights and Variations in the Number of Clusters

Fig. 8. shows the retrieval accuracy which inclined to increase until it reaches a variation of the number of clusters of 10 and inclined to decrease after the number of clusters is greater than 10.

3.2.3 Computing Time

The computation time that takes to identify an image in the medicinal plant leaf image database is influenced by the size of the image database and the database management techniques. The average computation time for the identification process in the image database before clustering is with variations in the number of clusters shown in Fig. 9.

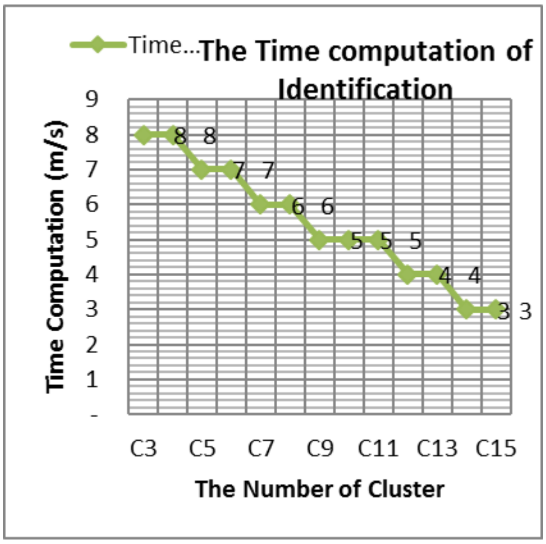


Fig. 9. Time Computation of Retrieval

Fig. 9 explains that the more the number of clusters, so computing time is getting faster. The highest number of time computation rates is owned by clusters *C3* and *C4*. *C5* to *C15* has a lower total time computation rate.

5 CONCLUSION

The level of accuracy in the identification process of medicinal plant leaves is influenced by several variables including image acquisition, preprocessing, feature extraction methods and identification techniques in the database. This study proves that the use of a different feature weighting scheme for each normalized feature value and using a variation in the number of clusters will have an effect on increasing the accuracy and speed of the medicinal plant identification process. The highest level of retrieval accuracy resulted in the variation of the number of clusters of 10 with a weighting variation of *W4* there are 60% shape features and 40% color features.

From this study it was found that the shape features in the image of medicinal plant leaves have a more dominant factor to determine the level of similarity in the identification process of medicinal plants. As for the color feature as a complementary feature. The average computation time required for retrieval is 5 milli-seconds.

REFERENCES

[1] A. Vadivel, A.K. Majumdar, and S. Shamik, "Characteristics Of Weighted Feature Vector In Content-Based Image Retrieval Applications," IEEE, 2004.
[2] R.C. Gonzales and R.E. Woods, "Digital Image Processing ThirdEdition. Pearson Prentice Hall," New Jersey, 2008.
[3] K.R. Castleman, "Digital Image Processing," Prentice Hall Inc., New Jersey, 1996.
[4] Jumi and A. Harjoko, "Image Similarity Analysis Based on Shape, Color and Texture Feature of Asset Image," International Conference on Computer Science Electronics and Instrumentation, Yogyakarta, Indonesia, 2012.
[5] J. Jumi, A. Zaenuddin and T. Mulyono, "Model for identification of rice type using combination of shape and color features," International Conference on Innovation in Science and Technology (ICIST), 2019.
[6] S. Azzam, L.A.D. Abdel, Q. Mohammad, A.S. Maryam, A.J. Rawa'a, and K. Ola, "Image Clustering using Color, Texture and

Shape Features," KSII Transactions on Internet and Information Systems, 5, 1, 211-227, 2011.
[7] H. Xing-yi, L. Jian, and J. Song, "Study on identification of rice varieties using computer vision [J]," Journal of Jiangsu University (National Science Edition), 2.003, 2004.
[8] K. Harpreet and S. Baljit. "Classification and grading rice using multi-class SVM," International Journal of Scientific and Research Publications, 3, 4, pp. 1–5, 2013.
[9] L. Fu, J. Duan, X. Zou, G. Lin, S. Song, B. Ji, Z. Yang, "Banana detection based on color and texture features in the natural environment," Computers and Electronics in Agriculture, 167, 105057, 2019.
[10] L. Vijaya, Balasubramanian, Mohan, Vasudev, "Kernel-based PSO and FRVM: An automatic plant leaf type detection using texture, shape, and color features," Computers and Electronics in Agriculture, 125, 99–112, 2016.
[11] T. Acharya, A.K. Ray, "Image Processing Principle and Applications, John Willey & Sons," USA, 2005.
[12] H. Hamdani, A. Septiarini, A. Sunyoto, S. Suyanto, and F. Utaminingrum, "Detection of oil palm leaf disease based on color histogram and supervised classifier," Optik, 245, 167753, 2021.
[13] A. Susilo, "Web Image Retrieval untuk Identifikasi bunga dengan Pengelompokan Content Warna," Institut Teknologi Sepuluh November, Surabaya, 2007.
[14] J. Chaki, N. Dey, L. Moraru, F. Shi, "Fragmented plant leaf recognition: Bag-of-features, fuzzy-color and edge-texture histogram descriptors with multi-layer perceptron," Optik, 181, 639–650, 2019.
[15] A. Susilo, "Web Image Retrieval untuk Identifikasi bunga dengan Pengelompokan Content Warna," Proceeding IES, Surabaya, 2006.
[16] Xiang Y. H, Yu-Jin B., and Dong H., 2003,Image Retrieval Based on WeigthedTexture Features using DCT Coefficients of JPEG Images, International Conference Information and Computer Security (ICICS), Singapore.