ISSTIN 2012

PROCEEDING

INTERNATIONAL SEMINAR ON SCIENCE AND TECHNOLOGY INNOVATIONS 2012

Green Technology Innovations for A Sustainable Society

University of Al Azhar Indonesia
2-4 October 2012

Organized by:
Faculty of Science and Technology
University of Al Azhar Indonesia

UAI
Universitas Al Azhar Indonesia

Sponsored by:

Islamic Development Bank

YAYASAN AGUNG PODOMORO LAND

JAYA ANCOL

REPUBLIKA

SOLUSI 247

Jaktv

ISBN:
978-602-95064-5-7

Published by:
UAI Press

APPLICATION OF IMAGE RETRIEVAL USING FRACTAL DIMENSION TO IDENTIFY MEDICINAL PLANT

Prihastuti Harsani[1], Iyan Mulyana[1], Prasetyorini[2]

[1]Computer Science Department Faculty of Mathematics and Natural Science
Pakuan University
[2]Biology Department Faculty of Mathematics and Natural Science
Pakuan University

E-mail: utik_harsani@yahoo.com, iyandelon@yahoo.com, prasetyorini@yahoo.com

Abstract - **Image-based search system to overcome the shortcomings of the text-based search system. Through image-based search system, said query keyword as the form of images so that the validity of the search results are expected to be larger. Method of fractal dimension is one of the methods for extracting the traits of leaves in plants, especially medicinal plants. Similarity between images as keywords and images contained in the specified database using Euclidean methods. Tests performed on the image of the leaves of medicinal plants which is the result of preprocessing. Preprocessing is done by changing the pixel size and coloration on the image to gray scale. The test results of image-based search system for the identification of medicinal plants produce 63% accuracy on the testing of medicinal plants with the same image and 68% with a different image of medicinal plants**

Keyowords *- image-based search technique, medicinal plants, Fractal dimension, Euclidian*

I. INTRODUCTION

Medicinal plants are plants that have medicinal properties and are used as drugs in the treatment or prevention of disease. Medicinal plants contain the active substance to treat a specific disease or resultant effect / synergy of the various substances that serve to treat [1]. To identify medicinal plants, a process that has to do is look at the book catalog of medicinal plants or by doing a search on medicinal plant information systems.

In general, the search technique is a text-based information system. It means the process require the keyword such as names and other code. When applied in the identification of medicinal plants, it is first necessary to know the name or code of the plant. The constraint is if the user does not know the name or code of the plant that will be identified.

This research developed information system with image-based search techniques. Search techniques based on similarity measurements between the image feature extraction leaves of medicinal plants with the query image feature extraction leaves of medicinal plants found on the image database. Feature extraction is performed using the method of fractal dimension while the similarity measurement using Euclidian Distance method. So that the identification of medicinal plants without needing to know the name and code.

II. BASIC THEORY

2.1.Fractal

Fractal comes from the Latin meaning is fractus (broken) or irregular. Fractal geometry is basically a simple one that can be broken up into several sections that have a shape like the previous form with a smaller size [2].

Fractal is characterized by fractal dimension in the form of fraction. One method used to calculate the

fractal dimension is box counting method (box counting) which can be expressed by [3]:

$$D = \frac{\log(N)}{\log(1/r)} \quad (1)$$

With
N= number of r-sized box that contains the object pixels
D= Fractal dimension
R= ratio

The steps box counting method is as follows:

a. The image is divided into boxes with size r. r value changed from 1 to 2^k, with k = 0, 1.2, ... and so on, and 2^k should not be larger than the size of the image. When the image size of 2m x 2m, then the value of k will not stop until m.
b. Count the number of boxes N (r) which contains parts of the object in the image. Value of N (r) depends on r.
c. Calculate the value of log (1 / r) and log (N).
d. Make a straight line using the log (1 / r) and log (N).
e. Calculate the slope of a straight line with equation 2 (Bruno et al, 2008). The value of this slope is the fractal dimension of the image of medicinal plants

$$\alpha = \frac{(n)\sum XY - \sum X \sum Y}{(n)\sum X^2 - (\sum X)^2} \quad (2)$$

With
α = number of data
X = value of log(1/r)
Y = value of log(n)

2.2 Euclidian distance

Euclidian distance is a method to measure the similarity of finding the smallest value or have the closest distance. The mathematical model can be expressed using Euclidean distance as the following equation [3]:

$$d_v = \sum_{j=1}^{m} ||F_{train\ to\ j,w} - F_{uji,w}|| \quad (3)$$

$$d_v = \sqrt{\sum_{j=1}^{m} (F_{train\ to\ j,w} - F_{uji,w})} \quad (4)$$

w is the characteristic index, the value of $Ft_{rain\ to\ j,\ w}$ is a feature matrix generated by the training data to j, with j ∈ 1 ... m and m = k * s. $F_{uji,w}$ is a row matrix. Similarity measuremen obtained are of minimal value of d_v or written using equation d = min (d_v)

III. METHODOLOGY

Stages of research carried out consists of the following steps (Figure 1). Image acquisition and preprocessing is preliminary preparation of the image to be extracted using fractal dimension. Data is stored in database includes general information on medicinal plants, medicinal plants and the benefits of the leaf image.

In the preprocessing stage had performed two sequential processes, namely: changes in image size and color. Image size is chaged into 16x16 pixel and color is converted to gray scale. The color change is not a thing that can affect the level of validity, but with a gray level, computational processes are expected to be faster when compared with the RGB or CYMK color image processing.

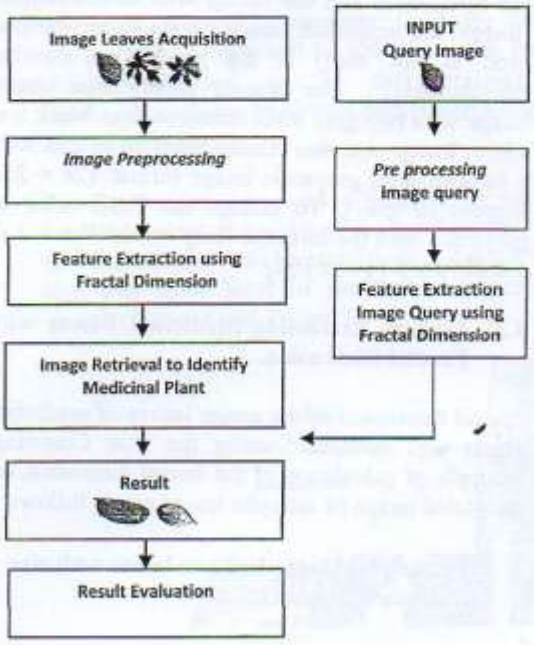

Figure 1. Steps of Research

IV. RESULT AND DISCUSSION

4.1. Image acquisition of Medicinal Plants

Stages of image acquisition are done through taking image of medicinal plant leaves using the line sensor (digital camera) and the array sensor (scanner).
The amount of data on medicinal plants has been the acquired as many as 300 species of medicinal plants and comprising 10 images of each medicinal plant leaves.

Data were collected from several sources, namely: Ministry of Health, Literatur Study and Collection of Bogor Botanical Garden Bogor.

4.2. Image preprocessing

Several step performed in the image processing are : the image segmentation, make the image size to be equal, and change the shape of the image into grayscale. In the process of image segmentation, the foregorund and the background of the objects (image) are separated. Image segmentation method used in this study is the process of floating (thresholding). The process generates a binary image with two gray level values such as black and white. Image size was standardized to 16 x16 with a form of 8-bit grayscale image format. (28 = 256 degrees of gray). To change the RGB color to grayscale used the formula: Gray = 0.2989 x R + G + 0.5870 x 0.1140 x B[2].

4.3. Feature Extraction Medicinal Plants with Fractal Dimension.

Fractal dimension of the image leaves of medicinal plants was calculated using the Box Counting. Example of calculation of the fractal dimension for the global image of Jatropha leaves are as follows:

1. Dividing the image leaf into boxes with size r. As shown in Figure 2.

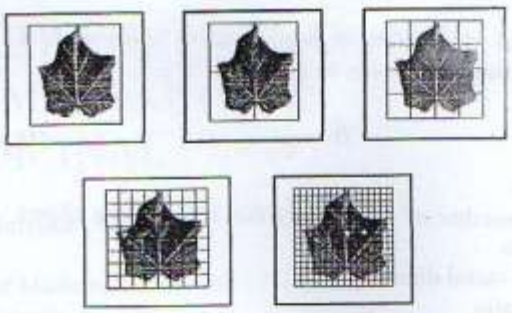

Figure 2. Divided image leaf into 16 boxes

2. Identifing r and N

r	1	1/2	1/4	1/8	1/16
N	1	4	14	47	148

3. Calculating value of Log (1/r) dan Log(N)

Log (1/r)	0	0,30103	0,60206	0,90309	1,20412
Log (N)	0	0,60206	1,14613	1,6721	2,17026

4. Displaying the Log(1/r) and Log (n) graph. As shown in Figure 3

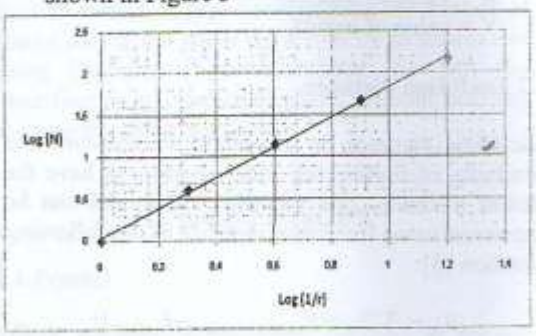

Figure 3. Chart comparison of the value of log (1 / r) and log (N).

5. Calculating gradien as fractal dimension, as shown in Table 1.

Tabel 1. Calculating X and Y

n	X Log(1/r)	Y Log(N)	XY	X2	Y2
1	0	0	0	0	0
2	0,30103	0,60206	0,18124	0,09062	0,36248
3	0,60206	114,613	0,69004	0,36248	131,361
4	0,90309	167,210	151,005	0,81557	279,591
5	120,412	217,026	261,326	144,990	471,004
∑	3,010	5,591	4,995	2,719	9,182

$$m = \frac{(n)\sum XY - \sum X.\sum Y}{(n)\sum X^2 - (\sum X)^2}$$

$$m = \frac{(5)4,995 - 3,010.5,591}{(5)2,719 - 3,010\ .3,010}$$

$m = 1,7973$ (nilai dimensi fraktal)

Image of leaves feature extracted generate 5 fractal dimension. The first value resulting from the extraction of the image as a whole (global). Four other value resulting from the extraction of the local region. Value of fractal dimension and shape of fragments ranging from one to two as shown in Figure 4.

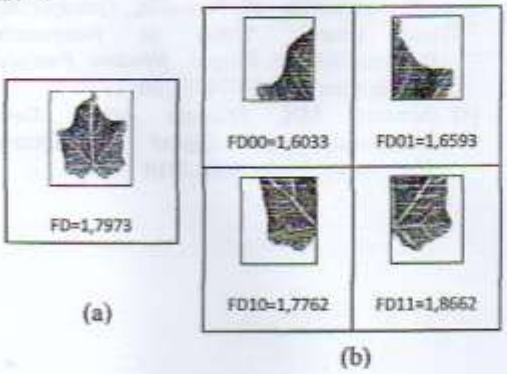

Figure 4. shows the application of extraction image feature using fractal dimesion.

Figure 5. Application for extraction feature image

4.4. Image Search-Based Applications for Medicinal Plant.

Application implementing image-based searching technique is used to measure similarity feature image extracted that will be identified with feature image extracted as a data training in database.Similarity level was determined using the euclidian method.

The application has been developed to identify medicinal plant. It consist of several function, such as : input image (will be identified) and result form that will show level of similarity between two image. (Figure 6 and Figure 7)

Figure 6. Application of medinical plant dientification

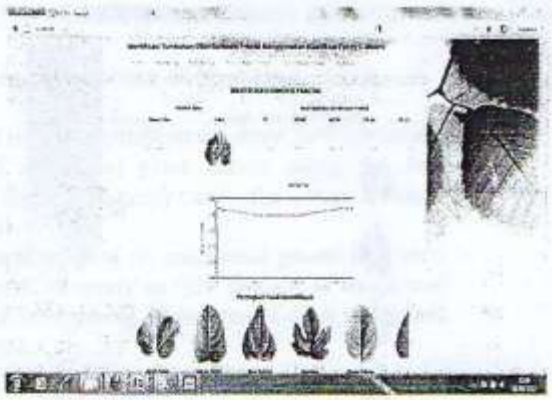

Figure 7. Search result

Applications developed to measure the similarity function of the image feature extraction leaves of medicinal plants to be identified with the image feature extraction leaf herbs contained in the database. Kemiripanya levels were measured using the Euclidian distance. Display an image-based search application for the identification of medicinal plants is shown in Figure 4 and Figure 5.

4.5. Testing

Testing are conducted to determine the success of the system. Trials carried out include, The first test conducted on medicinal plants query that has the same image with medicinal plants in the database. While the second test carried out on medicinal plants that have a query with a different image of medicinal plants in the database.

Validity of the system is represented using a matrix confusion. This value is measured between a medicinal plant in the training data and data testing. The accuracy obtained for the two consecutive tests was 63% for the same medicinal plants and 68% for different medicinal plants.

V. CONCLUSION

Based on research conducted conclusion can be obtained as follows:

1. The fractal dimension method and Euclidian distance can be implemented in developing Image-based search system for identification of medicinal plants.
2. The test results of image-based search system for the identification of medicinal plants produce 63% accuracy on the testing of medicinal plants with the same image and 68% with a different image of medicinal plants. The low level of accuracy is due to the extraction of the fractal dimension is more influenced by the pattern shape image of the medicinal plant leaves and there are several medicinal plants that have almost the same image.

REFERENCES

[1] Saifudin A, Rahayu V, Yuda H. *Standardisasi Bahan Obat Alam*, Yogyakarta: Graha ilmu, 2011
[2] Bruno OM, Oliveira , Falvo, Castro. Fractal dimension applied to plant identification. Sciencedirect 178:2722-2733, 2008
[3] Putra D, *Pengolahan Citra Digital*, Yogyakarta.Andi, 2010
[4] Jianxin Xu, Wang Hua, Fang Hui. Characterization of Periodeic, Quasiperiodeic and Chaotic States in Nonpremixed Biodiesel/Air Jet Flames. *Hindawi Publishing Corporation* 2011 (861436), 2011
[5] Purnomo MH, Muntasa Arief.. *Konsep Pengolahan Citra Digital dan Ekstraksi* .Yogyakarta: Graha ilmu, 2010