

Development of Feature Extraction on Leaf Image for Medicinal Plants Identification

Trinugi Wira Harjanti, Sarifuddin Madenda

Abstract: *The leaf image identification process depends on the feature extraction results. Each medicinal plant has different shapes and patterns of leaf venation. But for one type of medicinal plants have the same pattern of venation shape and pattern even though the size is different. One of the methods for extraction of leaf image form characteristics is by fractal-based feature extraction. Through fractal can be calculated the value of leaf dimensions and searched parts of leaves that have similarities between one part with other parts. As for the method of extracting the characteristics of venation pattern using B-Spline method. Benefits of research conducted is to help people identifying the types of medicinal plants found, knowing the benefits and ways of brewing. While the research contribution is prototype software application based on information technology that can be used by the people through mobile phones for the identification of medicinal plants. To identify or match the results of feature extraction on the leaf found whether included in the medicinal plant, conducted by Euclidean Distance method. In the experiments we used 1100 data consist of 55 variety of medicinal plants for each 20 samples. The experimental result show that the accuracy of identification using of fractal and b-spline is 85.30%.*

Index Terms: *Feature extraction, Medicinal plant identification, Fractal, B-spline, Euclidean distance*

I. INTRODUCTION

The support of the Government of the Republic of Indonesia in traditional health services can be seen from upstream, at the Ministry of Agriculture, there is the Balai Penelitian Tanaman Rempah dan Obat (Balitro). Balitro is a structural unit of technical implementation in the field of research and development under the Plantation Research and Development Center, Agricultural Research and Development Agency of the Ministry of Agriculture of the Republic of Indonesia. Meanwhile, from the downstream side, at the Ministry of Health, there is a Balai Besar Penelitian dan Pengembangan Tanaman Obat dan Obat Tradisional (B2P2TO-TO) is a Technical Implementation Unit of Health Research and Development Agency of the Ministry of Health of the Republic of Indonesia. Data Riset Kesehatan Dasar (Riskesdas) 2013, a national health research conducted by the Health Research and Development Agency of the Ministry of Health, shows that 30.4% of households in Indonesia use traditional health services, 77.8% of households use health services traditional skills without tools, and 49.0% of households utilizing potions. Another national scale research also conducted by the Health Research and Development Agency of the Ministry of Health is the Riset Tumbuhan Obat dan Jamu I (Ristoja) in 2012. The

study succeeded in obtaining data of 1.889 species of medicinal plants, 15.671 herbs for health, and 1,183 traditional healers from 20 % ethnicity (209 out of a total of 1.128 ethnic) Non-Javanese and Balinese [1]. With the diversity of medicinal plants making identification difficult, the ability to accurately identify medicinal plants is an essential requirement for experts and people involved in the world of medicinal plants [2].

The identification process depends on the feature extraction results. One part of the medicinal plant that can be used for the extraction is the leaves. Each medicinal plant has a different shape and leaf texture. But for one type of medicinal plants have the same pattern of venation shape and pattern even though the size is different. Pattern shape and pattern of venation of leaves can be used as a type of plant medicinal plants with each other. The founders used in this regard are characteristic of the shape, the pattern of the leaf venation and the combination of these two traits to produce more accurate image leaf recognition. One of the methods for extraction of leaf image form characteristics is by fractal-based feature extraction. Through fractal can be calculated the value of leaf dimensions and searched parts of leaves that have similarities between one part with other parts. This is because fractals have self similarity properties, which consist of parts that have similarities from one part to another [3]. There are two feature extraction approaches with the fractal method of fractal dimension and fractal code. The fractal dimension is based on the uniformity of the shape pattern that an image possesses. While the fractal code is based on the resemblance of the texture patterns on the self of an image [4].

Some studies using fractal method are used fractal dimensions in Iris Recognition System with classification using three methods: Bayes, Euclidean and K-nearest neighbor (K-NN) [5]. The accuracy rate in this study is 100% for all training data and 90% for different data. Mozaffari (2005) uses fractal codes for pattern recognition of Arabic characters and digits by classification using the Support Vector Machine (SVM) and Radial Basis Function (RBF) Neural Network. The results of this study showed that classification with SVM achieved 91.33% pattern recognition rate for the characters and 92.71% for the digits [6]. While the classification with RBF reaches 90.9% pattern recognition rate for characters and 91.70% for digits.

As for the extraction of leaf venation pattern, one method that can be used is using B-Spline model. The extraction process involves searching the structure by interpreting spatial information into a value. The main spatial venation information is represented in the form of b-spline S_m . With the help of S_m , continued search of side venation of leaf and finally represented also a list of b-spline Rahmadhani (2009)

Revised Manuscript Received on April 25, 2019.

Trinugi Wira Harjanti, Information System, STTI NIIT, 13770, Indonesia

Sarifuddin Madenda, Computer Science & Information Technology, Gunadarma University, 16424, Indonesia

used the b-spline model for the extraction of leaf venation patterns [7].

To identify or match the results of feature extraction on the leaf found whether included in the medicinal plant, conducted by Euclidean Distance method. Some studies using Euclidean Distance are Shabrina (2006) study on distance analysis using Euclidean Distance as the basis of measuring similarity of iris image [9]. From the results of research [9] Implementation of Euclidean Distance on Pattern Matching for Image Conversion to Text, it can be seen that the use of Euclidean Distance algorithm gives maximum result compared to Fast Fourier Transform algorithm, both in terms of text file result and from in terms of computing time [10].

II. METHODOLOGY

The stages of the research can be shown at Fig.(1).

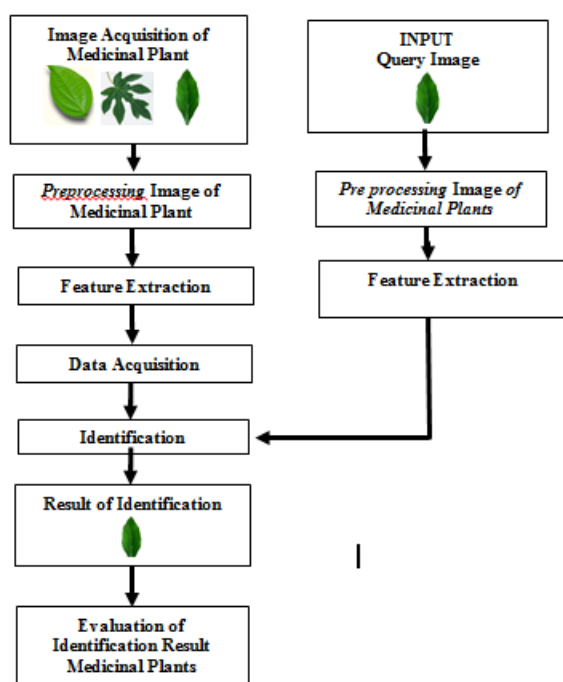


Fig. 1 Stages of the research

A. Image Acquisition of Medicinal Plants

At this stage the image of medicinal plants is accumulated with array sensor (sensor array) using 8 MP digital camera. The image data is stored in JPG format with size 3264x2448 pixels. The data used in this study is a collection of medicinal herbs in the garden of the Balai Penelitian Tanaman Rempah dan Obat (Balitro) Center for Plantation Research and Development, Agency for Agricultural Research and Development Ministry of Agriculture Republic of Indonesia, Bogor.

In this study the image of medicinal plant leaf used is 55 species and each type consists of 20 leaf image so that the total is 1100 data.

B. Preprocessing Image of Medicinal Plants

Implementing Image Segmentation

The first preprocessing step performed in the study was to segment the leaf image by removing the background or background from the leaf image.

Converting Image Color to Grayscale

The next preprocessing phase is to change the shape of the image into an 8 bit grayscale format (28 = 256 degrees gray). To change the color of RGB to grayscale used equation [8].

$$\text{Gray} = 0.2989 \times R + 0.5870 \times G + 0.1140 \times B$$

Uniform Image Size

The next preprocessing phase is to uniform the size of the image to 128x128 pixels. This is done because the resolution of the input image to be processed is too large, also to ensure that any area taken as an input on a leaf will produce the same shape and pattern of venation leaves and to overcome difficulties in taking the leaf dataset intact.

C. Feature Extraction

Feature Extraction Using Fractal Dimension Method

In this study, the feature extraction step was performed on the leaf image of the shape. The extraction of form characteristics to the leaf image is done by using Fractal method. The fractal method approach used for feature extraction is the fractal dimension.

Fractal dimension of medicinal leaf plant image is counted by using box counting method which is as the same as [10]:

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

with :

N = a number of measurement box r which filled with an object pixel

D = fractal dimension object

r = ratio

These are the box counting levels :

1. Dividing the leaf image into squares of size r
2. Identify the r and N value
3. Calculate the log value $(1/r)$ and $\log(N)$
4. Making a straight line by using value $D(s)$
5. Calculating the value of the slope of the line as the fractal dimension

Feature Extraction Using B-Spline Method

The extraction process involves searching the structure by interpreting spatial information into a value. The main spatial venation information is represented in the form of b-spline S_m . With the help of S_m , continued search of side venation of leaves and finally represented also a list of b-spline S .

For non-homogeneous and reflective illumination compensation, a high-pass $H = 1 - B$ filter is applied, where B is a binomial smoothing filter. Then initialization of starting point parameters and the direction of search structure interactively.

After initializing the parameter, the venation is searched separately. The main venation search is done first. The main venation points found to be the basis of a b-spline and obtained a continuous description of the main venation, called S_m .

The search for side venation is done by taking the sample area around the main venation by b-spline method. Sampling is done along the normal line to all b-splines found at each b-spline length measured in pixels. Image of adjacent image of main venation is obtained.

All the maximum values on parallel lines of the main venation are taken as possible starting points in the search for side venation structures. From each maximum position, the search for the structure is performed in the normal direction of the main venation obtained from S_m and converts the points found into the b-spline interpretation S_s .

D. Data Acquisition Phase

The national scale research that has been done by the Health Research and Development Agency of the Ministry of Health is the RisetTumbuhanObatdanJamu I (Ristoja) in 2012. The study succeeded in obtaining data of 1.889 species of medicinal plants, 15.671 herbs for health, and 1.183 traditional healers from 20% ethnicity (209 out of a total of 1.128 ethnic) non-Javanese and Balinese. Based on these data, namely the name of medicinal plants, descriptions, efficacy and how the potion, then put into the database as the basis of information that can be used community.

E. Phase Identification of Medicinal Plants with Euclidean Distance

Identify the image of a tested medicinal plant or query to find out the name of the plant, its benefits and its potion. The way is by matching the characteristic extraction value of the test image with the extraction value contained in the database using the Euclidean Distance method.

Identification is done on 1100 training data. The steps are as follows:

- Data inputs are $n \times m$ size matrices (n = number of data samples, m = data attributes which are the fractal dimension values and fractal value codes in the first experiment, co-occurrence matrix values in the second experiment, combined with fractal dimension values, values fractal codes, and co-occurrence matrix values in the third experiment.
- Compare the minimum distance of test image (testing), with database image training (training). Euclidean distance of two vectors x and y is calculated by Equation:

$$D = (|x_0 - x_1|^a + |y_0 - y_1|^a)^{1/a}$$

The smaller the value $d(x, y)$, the more likely the two vectors are matched / compared. Conversely, the greater the value $d(x, y)$ the more different the two vectors are matched.

F. Evaluation of Identification Result Medicinal Plants

In this study there are three tests conducted to determine the success of image identification of medicinal plants, namely:

- Testing identification of the same leaf image between test data and training data using fractal dimension.

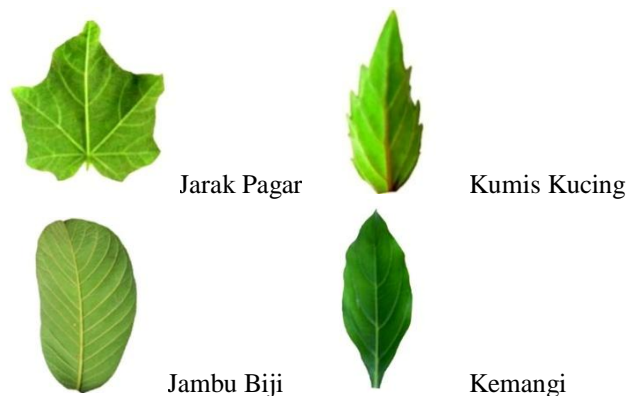
- Uniform leaf image identification test between test data and training data using fractal dimension.
- Testing identification of the same leaf image between test data and training data using b-spline method.
- Uniform leaf image identification test between test data and training data using b-spline method.
- Testing identification of the same leaf image between the test data and training data using a combination of fractal dimension and b-spline method.
- Uniform leaf image identification testing between test data and training data using combined fractal dimension and b-spline method.

To calculate the level of accuracy of the test results used confusion matrix between real data and predicted data from leaf image of medicinal plants.

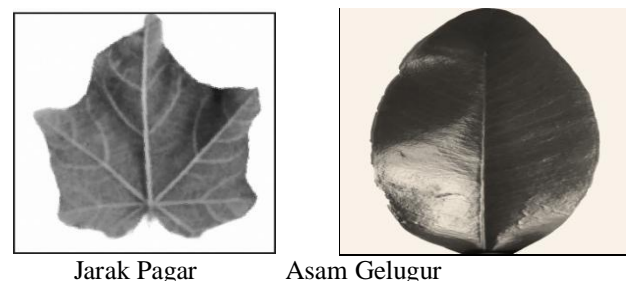
III. RESULTS AND DISCUSSION

A. Preprocessing Image Result of Medicinal Plants

The first preprocessing step performed in the study was to segment the leaf image by removing the background or background from the leaf image. Here is an example of leaf image segmentation that has been done:



The next preprocessing phase is to change the shape of the image into an 8 bit grayscale format (28 = 256 degrees gray). Here is an example of an image that has been converted into a grayscale format:

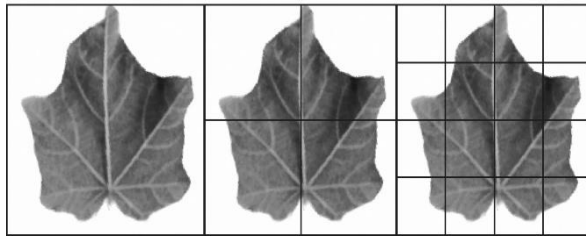


B. Feature Extraction Result with Fractal Dimension

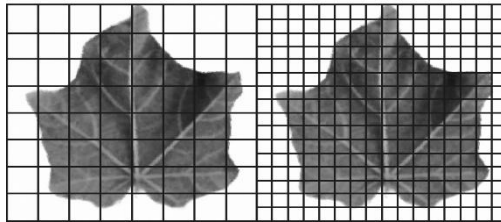
The fractal dimension of plant leaf image is calculated using Box Counting method. Examples of global fractal dimensional calculations for Jatropha leaves images are as follows:



- Dividing the leaf image into squares of size r . As shown in Fig. 2.



$r=1, N=1$ $r=2, N=4$ $r=4, N=14$



$r=8, N=47$ $r=16, N=148$

Fig. 2 Leaf images divided into squares of size r .

- Identify the r and N Value

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

- Calculate the log value ($1/r$) and $\log(N)$

Log ($1/r$)	0	0.30103	0.60206	0.90309	1.20412
Log (N)	0	0.60206	1.14613	1.6721	2.17026

- Displays a graph of $\log(1/r)$ and $\log(N)$ value. As shown in Fig. 3.

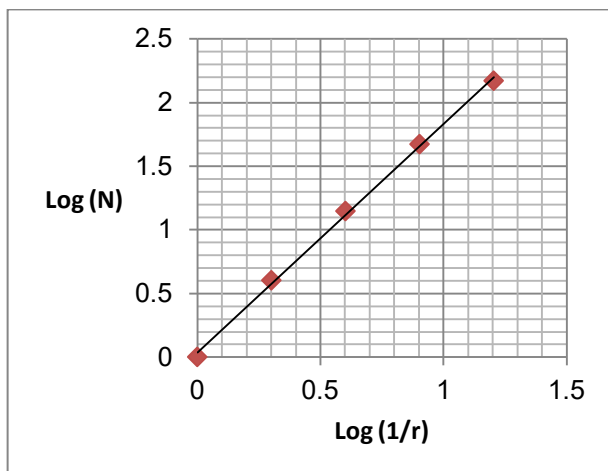


Fig. 3. Graph value of the log value ($1/r$) and $\log(N)$

- Calculating the value of the slope of the line as the fractal dimension as shown in Table 1.

Table 1 Calculation of values of x and y

N	X	Y	XY	X ²	Y ²
	Log ($1/r$)	Log(N)			
1	0	0	0	0	0
2	0.30103	0.60206	0.18124	0.09062	0.36248
3	0.60206	1.14613	0.69004	0.36248	1.31361
4	0.90309	1.67210	1.51005	0.81557	2.79591
5	1.20412	2.17026	2.1326	1.44990	4.71004
Σ	3.010	5.591	4.995	2.719	9.182

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

$$\alpha = \frac{(5) 4.995 - 3.010 \cdot 5.591}{(5) 2.719 - 3.010 \cdot 3.010}$$

$$\alpha = 1.7973 \text{ (fractal dimension value)}$$

C. Data Acquisition Phase

The national scale research that has been done by the Health Research and Development Agency of the Ministry of Health is the Riset Tumbuhan Obat dan Jamu I (Ristoja) in 2012. The study succeeded in obtaining data of 1.889 species of medicinal plants, 15.671 herbs for health and 1.183 traditional healers from 20% ethnicity (209 out of a total of 1.128 ethnic) non-Javanese and Balinese [1]. Here is an example of the resulting data:

ABELMOSCHUS MOSCHATUS MEDIK

Botani
Sipom
Klasifikasi
Divisi
Sub divisi
Kelas
Bangsa
Suku
Marga
Jenis
Nama umum/dagang
Nama daerah
Sulawesi
Jawa
Maluku
Deskripsi
Habitus
Batang
Daun
Bunga
Buah
Biji
Akar
Khasiat

Hibiscus abelmoschus L.
Spermatophyta
Angiospermae
Dicotyledonae
Malvaceae
Abelmoschus
Abelmoschus moschatus Medik
Waron
Kapas sedeli (Lampung)
Kakapasan (Sunda) Waron (Jawa) Kastore bukal (Madura)
Kashuri (Ternate)
Semak, semutim, tinggi ± 3,5 m.
Berkasus, bulat, tegak, percabangan monopodial.
Tangkai, pertapi lima, berlekuk, bercangap atau terbagi lima, pangkal bentuk jantung, ujung (snop) tepi rata, panjang 6-22 cm, lebar 5-20 cm, pertulangan menyari, tangkai pangkal 5-10 cm, hijau.
Tangkai, bentuk lonjong, di kelaki daun, kelopak 2-3 cm, segi tiga, berbulu, ujung bertaji lima, hijau, benang sari bentuk tabung, kepala sari lepas, kuning, malikola lima, pangkalnya merah, panjang 3,5 - 10 cm, kuning.
Pistak, bulat telur, noksi lima, mancong, berbulu atau berambut tipis sikat, berlekuk, keras, coklat kehijauan.
Bentuk ginjal, pipih, berlekuk, keras, coklat kehijauan.
Tunggai, bulat, bercabang, puth kekuningan.

Dan Abelmoschus moschatus berhasiat sebagai obat batuk, bijnya merupakan bahan baku kosmetik.
Untuk obat batuk dipakai ± 2 gram serbuk daun Abelmoschus moschatus, diseduh dengan 1/2 gelas air matang panas, setelah dingin disaring. Hasil saringan diminum setiap dua kali 1-4 gelas pagi dan sore.

Kandungan kimia

Biji, daun dan akar Abelmoschus moschatus mengandung saponin dan flavonoida, di samping itu daunnya juga mengandung tanin dan bijnya juga mengandung tanin dan bijnya juga mengandung minyak atsiri.



Gambar 1. Abelmoschus moschatus Medik.

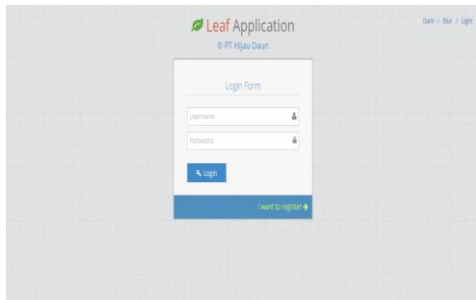
Fig. 4. data on the benefits of medicinal plants

Based on these data, namely the name of medicinal plants, descriptions, efficacy and how the potion, then put into the database as the basis of information that can be used community.

D. Web Application Display Identification of Medicinal Plants (Admin)

- Login





login user will get notification that there is new data to be input into the database.

E. Display of Mobile Application Identification of Medicinal Plants

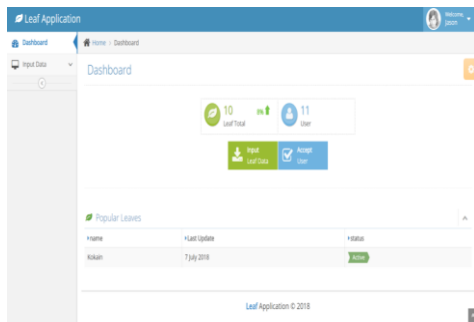
Consists of 3 main menu (Scanning, Plants, Medicinal Plant):

- Scanning



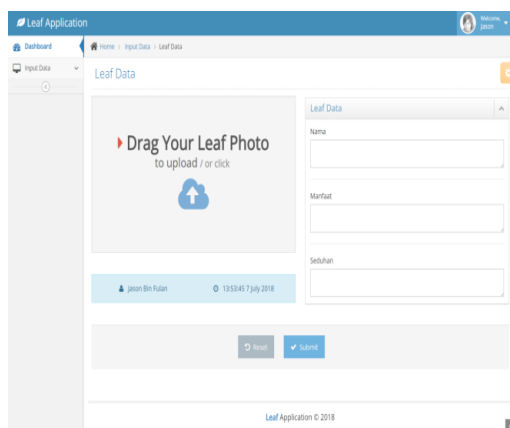
This menu serves to photograph the leaves or take a photo / image leaves from local drive (mobile phone), which will then be done the process of identification of the photo / image of the leaf.

- Dashboard



On the dashboard page presented the leaf view data of the most in search by mobile users. There are also total leaves and total users, so users can easily know the data. There is also a shortcut to add new leaf data, or receive a registration request from a new user.

- Leaf Image input



On the data input page, the user can drag the photo of the leaf to be uploaded. Once uploaded, you can enter data from the leaf. After the user ensures the data and photo leaves are correct, then the user can click submit to enter data into the database or reset if you want to empty the entire form.

- Case Data Not Found by User Mobile

When a mobile user scans a leaf and data is not found in the database, the user can send feedback to the team via the mobile application. After the mobile user sends feedback, the

- List of plants



This menu displays a list of medicinal plants already recorded in the database (Medicinal Plant Identification App) End Users can search for medicinal plants by typing the name of the medicinal plant in the search field where the search results will be displayed and can be clicked / selected after it will be raised short information from medicinal plants that have been in the click / select it

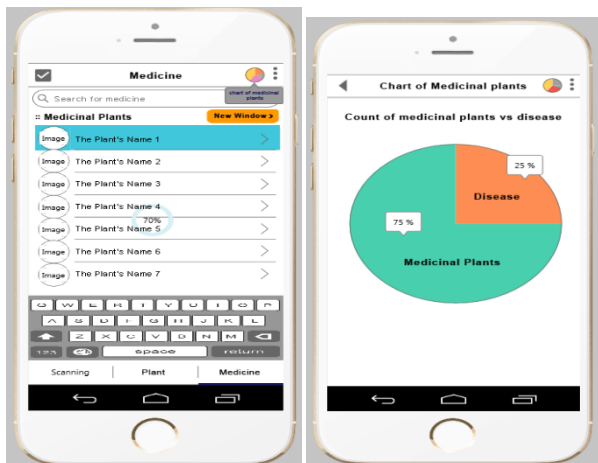
- Medicinal Plant

This menu displays a list of medicinal plants in certain diseases, the list is data recorded in the database (App Medicinal Plant Identification). To know the medicinal plants of a particular disease End Users can type the name of the search field disease.



The graph icon is a link to the medicinal plants chart vs. disease

New Window > is a link to a new page that displays a list of medicinal plant search results from a particular disease



- Brief Information of Plants

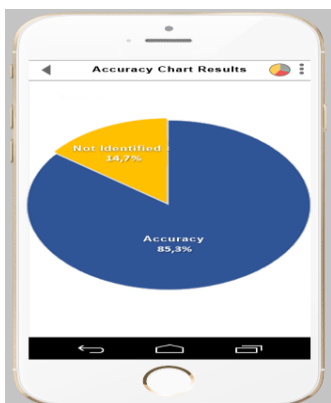
Displays short information of plants from the process of scanning, list of plants, list of medicinal plants against disease. Information displayed include photos, crop names, plant Latin names and plant properties.



button has a function to call and display complete information (pdf) of the plant.

- Accuracy value

The experimental result show that the accuracy of identification using of fractal and b-spline is 85.30%



study resulted 85.30% accuracy, making it easier for users to know the types of medicinal plants they meet accurately and quickly. In addition to the medicinal plant species, the system is also expected to provide information on the plant covering the properties, the parts used for the medicinal, its characteristic, chemical content, and its usefulness so as to help the people in developing traditional medicine with the resources around them.

REFERENCES

1. Y. Aditama. *Jamu dan Kesehatan*. Jakarta: Lembaga Penerbitan Balitbangkes. Badan Litbang Kesehatan Kementerian Kesehatan Republik Indonesia, 2014.
2. I. Mulyana Y. Herdiyeni and S. H. Wijaya. "Identification of Medical Plant Based on Fractal by Using Clustering Fuzzy C-Means." In Prosiding ICIBA, 2013.
3. Mandelbort, The Fractal of Nature, Berlin: Springer Verlag, 1982.
4. M. F. Barnesley, R. L. Devaney, P. Mandelbort, D. Saup and R. F. Voss. The Science of Fractal Images. Berlin: Springer verlag, 1988.
5. M. P. S. Chandra, S. Reedy and B. Ramesh. "Iris Recognition System Using Fractal Dimension of Haar Patterns." Internasional Journal of Signal Processing, vol. 2, pp. 75-81, 2009.
6. S. Mozaffari, K. Faez and H. R. Kanan. "Performance Evaluation of Fractal Feature in recognition of Postal Code Using an RBF neural Network and SVM Classifier." In IAPR Conference on Machine Vision Applications, pp. 562-565, 2005.
7. M. Ramadhani. Extraction of Form Features and Leaf Image Vanation by Fourier and B-Spline Modeling. Bachelor Thesis, Institut Pertanian Bogor, 2009.
8. Jianxin Xu, Wang Hua, Fang Hui. 2011. Characterization of Periodic, Quasiperiodic and Chaotic States in Nonpremixed Biodiesel/Air Jet Flames. Hindawi Publishing Corporation 2011 (861436)
9. K. Shabrina, Distance Analysis Using Euclidean Distance as a Basis for Measuring the Similarity of Iris Eyes, 2006
10. F. Samopa and Y. Yulianawati. "Applacation of Euclidean Distance on Pattern Matching for Image Conversion to Text." JUTI: Jurnal Ilmiah Teknologi Informasi, vol. 1, no. 1, pp. 32-42, 2002.

IV. CONCLUSION

The results of this study are the creation of a medicinal plant identification system that can be implemented on mobile devices so that it is easy to use by anyone and anywhere. The combination of shape and venation extraction of leaf medicinal plants using fractal and b-spline in this