

# Identification of Plant Types by Leaf Textures Based on the Backpropagation Neural Network

Taufik Hidayat, Asyarah Ramadona Nilawati

Department of Informatics, Gunadarma University, Indonesia

---

## Article Info

### Article history:

Received Sep 30, 2017

Revised Feb 21, 2018

Accepted Jul 22, 2018

---

### Keyword:

Feature

Leaf

Perimeter

Ratio

Roundness

Slenderness

---

## ABSTRACT

The number of species of plants or flora in Indonesia is abundant. The wealth of Indonesia's flora species is not to be doubted. Almost every region in Indonesia has one or some distinctive plant(s) which may not exist in other countries. In enhancing the potential diversity of tropical plant resources, good management and utilization of biodiversity is required. Based on such diversity, plant classification becomes a challenge to do. The most common way to recognize between one plant and another is to identify the leaf of each plant. Leaf-based classification is an alternative and the most effective way to do because leaves will exist all the time, while fruits and flowers may only exist at any given time. In this study, the researchers will identify plants based on the textures of the leaves. Leaf feature extraction is done by calculating the area value, perimeter, and additional features of leaf images such as shape roundness and slenderness. The results of the extraction will then be selected for training by using the backpropagation neural network. The result of the training (the formation of the training set) will be calculated to produce the value of recognition accuracy with which the feature value of the dataset of the leaf images is then to be matched. The result of the identification of plant species based on leaf texture characteristics is expected to accelerate the process of plant classification based on the characteristics of the leaves.

*Copyright © 2018 Institute of Advanced Engineering and Science.  
All rights reserved.*

---

## Corresponding Author:

Taufik Hidayat,

Department of Informatics, Gunadarma University,

Margonda Raya Street 100, Pondok Cina, Depok, Indonesia.

Email: thidayat@staff.gunadarma.ac.id

---

## 1. INTRODUCTION

Indonesia is one of the countries with high plant diversity. The number of species of plants or flora in Indonesia is abundant. The wealth of the species of Indonesia's flora is unquestionable [1]. Almost every region in Indonesia has one or some distinctive plant(s) which may not exist in other countries. Not only the diversity but some types of plants in Indonesia have many benefits for health. In enhancing the potential diversity of tropical plant resources, good management and utilization of biodiversity is required. Geranium plants in Indonesia, as an example, have some species. Geranium flower plants are also known as herbal plants called tapak dara. The geranium flower belongs to the plant producing the essential oil and is categorized as a family of Geraniaceae. This plant, in the Indonesia's herbal family, is known as tapak dara flower while in Latin it is named as *Palrgonium graveolens* [2]. The other plant such as jabon or in Latin is referred to as *Anthocephalus Cadamba* which grows wild in the forest and which has become the popular one of the alternative herbs in Indonesia in recent years. Currently, many Jabon plants are cultivated because of their advantages compared to other plants [3]. The plant belongs to the one that is very easy to grow in Indonesia, because the climate supports the perfect growth of this Geranium flower. It is estimated that there are 2 millions of plant species worldwide that have been recognized and 60% of them are in Indonesia; however, up to the present, the exact number of how many plant species having been grown in Indonesia

could not have been exactly determined. Currently there are only 8,000 species that have been identified. The amount is estimated only 20 percent of the total flora that exists in Indonesia[4].

Based on such diversity, plant classification becomes a challenge to do. The most common way to recognize between one plant and another is to identify the leaf of each plant. Leaf-based classification is an alternative and the most effective way to do because leaves will exist all the time, while fruits and flowers may only exist at any given time. Classification of fruit plants based on leaves can be done on the basis of the morphological characteristics of textures that can be observed or measured from the leaves or the images of the leaves[5]. Some researches related to the plant identification based on textures, morphology, and leaf colors have been done by the previous researchers. The research on classification based on color textures and leaf shapes was carried out by the researchers[6] by using the Probabilistic Neural Network with the supervised training and the Feed forward structure. Bayes' rule of the Kernel Fisher Discriminant Analysis was used to classify a number of leaf categories. Decision making was based on the result of calculating the distance between the probability density function of the characteristic vector based on the roundness and slenderness of the leaf images [6]. The research on identifying plants based on leaf shapes was carried out by the Researchers [7] using multilayer perceptron neural network (MLP). The researchers used perceptron with one weight layer which has only a linear function with the input of approximately 6 species out of 197 leaves with similar structures such as mango, sapota, guava, neem, and cotton. The result showed that MLP has a leaf classification accuracy value of 88.20% [7]. Identifikasi terhadap jenis tanaman Adenium dilakukan [8] menggunakan metode Learning Vector Quantization. This method is used by the researchers to classify adenium plants where each output will represent a class or category of Adenium. In this method there may be multiple outputs for each class. The weight vector for an output unit is usually a reference to the class in which the unit is located. The learning method in this study will classify input vectors, classes and spacing between input vectors. If two input vectors have approximately equal spacing, then both input vectors will be placed into the same class [8]. Another research in plant identification based on leaf characteristics was also conducted by the researcher using the Extreme Learning Machine (ELM). ELM is a single-layer feedforward neural network or usually abbreviated as SLFNs. There are many types of feedforward artificial neural networks. The learning process of ELM is much slower than expected because all parameters are given manually and iterative tuning is required on each parameter [9].

In this study the researcher will identify plant types by their leaf textures. Leaf feature extraction is done by calculating the area value, perimeter, and additional features of the leaf images such as the roundness and slimness of the leaf shapes. The results of the extraction will then be selected for training using the backpropagation neural network. The training result (the formation of the training set) will be the calculation of the value of recognition accuracy with which the feature value of the dataset of the leaf images is then to be matched.

## 2. RESEARCH METHOD

### 2.1. Data Sets

In this study, the treated images are the leaf digital images obtained from the UCI Machine Learning Repository [10]. The data of the plants' leaf images have the size of 2322x4128 pixels with the jpg-formatted RGB images. The processed images are taken from the leaf dataset comprising 32 classes of plants as what can be seen in Figure 1.



Figure 1. The example of the leaf dataset [10]

Each class has species with varying amounts of data per class. The total number of images used is 1605 images of plant leaves. For example, in the Geranium class, Figure 1 has 15 variations of leaves as what can be seen in Figure 2.

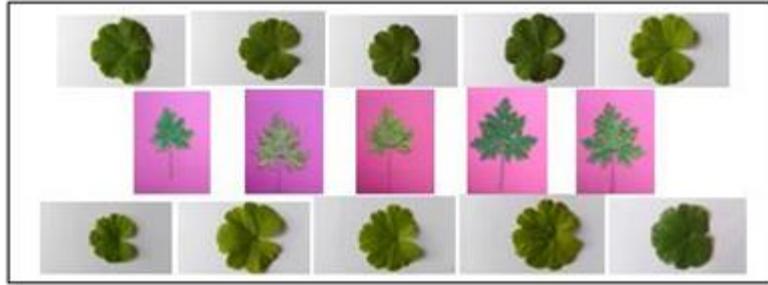


Figure 2. The example of the leaf dataset [10]

## 2.2. Leaf Segmentation

Leaf image segmentation process is done to separate the foreground from the background of each leaf. The leaf image segmentation process conducted by the researchers is done through the following stages:

- a. Execute the blue channel extraction of each leaf image. The blue channel is used as the main color because the blue color has the highest intensity of the other two colors—red and green—of each image within the same type of the RGB colors as what can be seen in Figure 3.

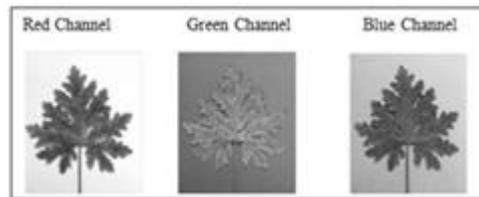


Figure 3. The channels of the leaf imagery

Here is the mathematical formula for finding the blue channels [11]:

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

Formula (1) is a green color channel while R, G and B are sequentially Red, Green, and Blue. Through the blue channel, the details of the image can be seen clearly and thoroughly while the use of the red channel will only display image restrictions, and through the green channel the image can only be partly seen and there is a lot of noise. In figure 3, it can be observed that leaf morphology appears most contrasting in the blue channel compared to the red and green channels in the RGB image.

- b. Perform the binarization process for the blue channel leaf images. The input is the original image and the output is the image resulted from the binary process. This binarization can be performed using formula 2 [12] :

$$g(x, y) = \begin{cases} 0 \rightarrow f(x, y) \geq T \\ 1 \rightarrow f(x, y) \leq T \end{cases} \quad (2)$$

The object extraction from the background is to select the threshold value T (T represents the pixel mapping value) that separates the two modes (0 and 1). Afterwards, for any point (x, y) that satisfies  $f(x, y) > T$  is called the point of the object, otherwise called the background point [12]. The image resulted from the binary process can be seen in Figure 4.

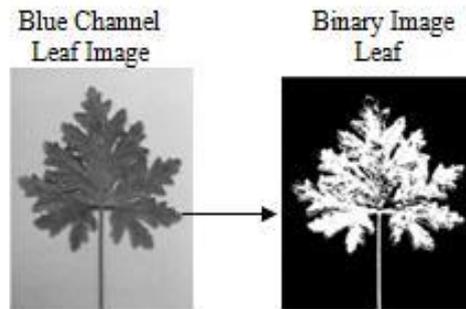


Figure 4. The leaf image binarization process

- c. Do a closing operation on the binary image to remove the black pixels inside the leaf object by enlarging the outer boundary of the foreground object and also close the small hole located in the middle of the object with the formula [12]:

$$A \bullet S = (A \oplus S) \otimes S \quad (3)$$

In the closing operation, the researchers use disc-shaped structuring elements to adjust the shape of the leaf image. The result of the closing operation can be seen in Figure 5.

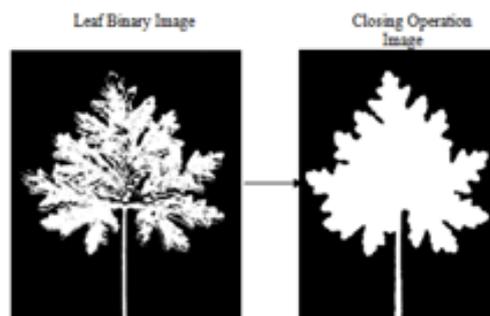


Figure 5. The closing operation of the leaf binary image

### 2.3. Leaf Feature Extraction

The previous research only used the texture feature as the one image feature in identifying the type of a plant. The study of the feature characteristic extraction towards the Ethiopia coffee plant disease was done with HSV color space where the features of the coffee leaves had different color variations [14]. The previous research did not use the shape feature which can visually show that a very different plant has a different leaf shape from other leaves. In this study, the researchers proposed the feature characteristics in identifying plant types based on their leaves by looking at the geometric shape of the leaf object- the roundness or the slenderness- by calculating the area of the leaf. A simple way to calculate the area of a leaf object is by counting the number of pixels on the object. The leaf feature extraction is based on the measurement using the object geometry approach which includes:

- a. The area value which is the number of pixels pertaining in the segmented image region of the leaf.
- b. The perimeter is a circumference that expresses the length of the surrounding edge of the leaf image object as what can be seen in Figure 6.

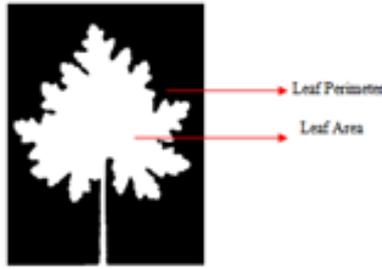


Figure 6. The leaf image perimeter

The edge of the leaf object is processed using a chain code, so that the perimeter can be calculated using the formula [8]:

$$Perimeter = X_{even} + X_{odd} \tag{4}$$

where:  $X_{even}$  =Even Code

$X_{odd}$  =Odd Code

Figure 7 is an example of using a chain code in calculating the perimeter of the leaf image. The area of the leaf object is calculated by counting the number of pixels on the leaf object.

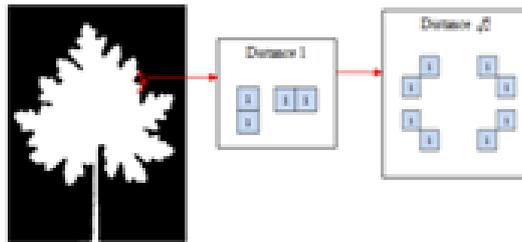


Figure 7. The chain code in determining the leaf image perimeter

c. The leaf image alternative feature which is used to determine the edge variation of the object on the leaf image is based on:

- 1) **The shape roundness.** It is a comparison between the object area and the perimeter square calculated by using the formula of roundness :

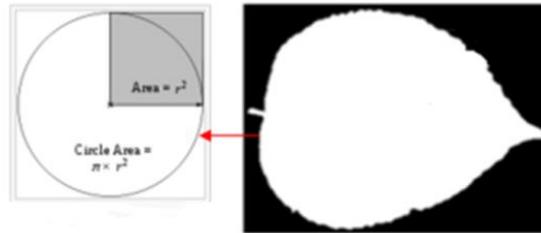
$$Ratio = \frac{Perimeter}{4\pi A} \tag{5}$$

Where [13] :

$$A = \pi r^2 \tag{6}$$

The area and the perimeter values which are the properties of this circle can be calculated on the leaf image regions which are extracted as the basic form of the roundness size [13]. The  $R$  ratio for a circle is  $4\pi$  which the minimum value for each region is. The ratio  $R$  value will produce values ranging from 0 to 1, where the value 0 is assumed that the leaf image object is circular as what can be seen in figure 8 of the component of the circle object.

The use of the length and the width features of a leaf object based on the ratio can show that a leaf shape with round or slender specification can be distinguished to facilitate the identification process of the plant type based on its leaf features



.Figure 8. The representation of the shape roundness of a leaf image

- 2) **Shape Slenderness.** It is the comparison form between the major axis length and the minor axis length as what can be seen in Figure 9.

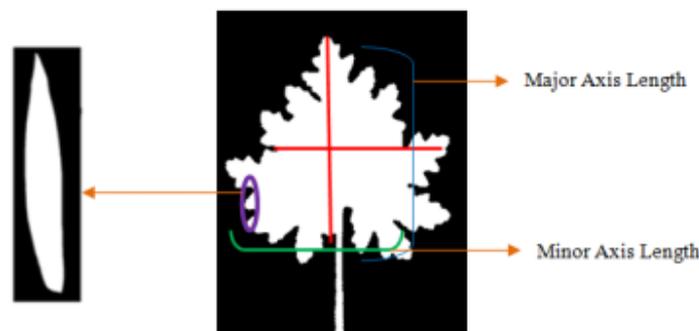


Figure 9. The major axis and minor axis length of a leaf image

The use of the length and the width features of a leaf object based on the ratio can show that a leaf shape with round or slender specification can be distinguished to facilitate the identification process of the plant type based on its leaf features.

#### 2.4. Leaf Image Training by the Backpropagation Neural Network

The training process performed to identify plants based on their leaf image textures requires a training set of parameters of the leaf image characteristics such as the feature of the area value, perimeter, and alternative features of the leaf image consisting of features of roundness and slenderness. Both of these features are the input to the neural network. As a trial, it is used as many as 1605 images of plant leaves taken from 32 classes of plants. The feature value is then used as the input to the training process using Levenberg-Marquardt as what can be seen in Figure 10.

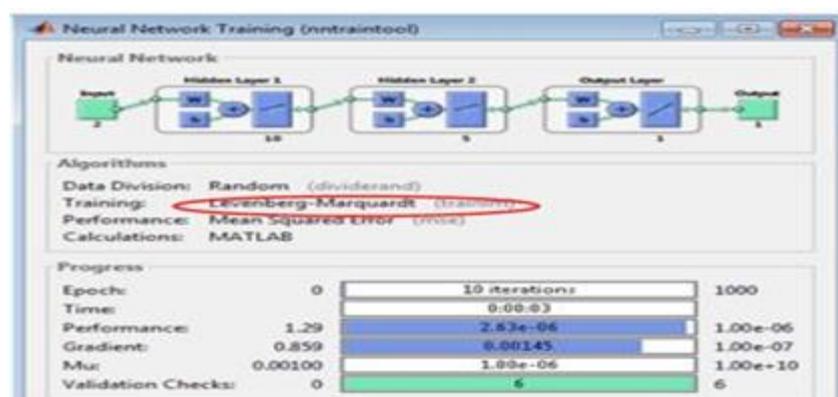


Figure 10. The scheme of the training process with neural network

**2.5. Training Target**

The initial stage of the training process is to establish the target matrix of 32 plant classes. For example, if there are 10 classes of plants, a target matrix is formed with a 10x10 matrix order as what can be seen in the formation of the target matrix in Figure 11.

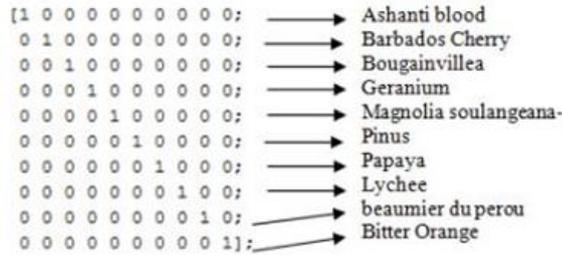


Figure 11. The illustration of the establishment of plant class targeted matrices

**2.6. Leaf Image Training Parameter**

The researchers used the backpropagation method with two hidden layers using the extraction of the leaf feature extraction of feature characteristics. The leaf image recognition process using the backpropagation Neural Network is done by determining some parameters as what can be seen in figure 12 of neural network scheme.

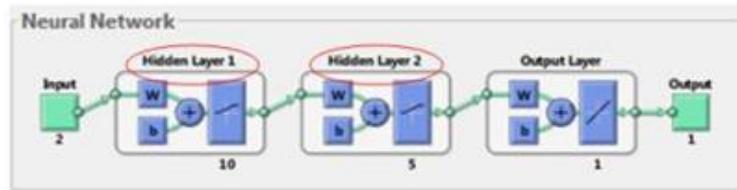


Figure 12. The neural network scheme

**2.7. Training Accuracy**

Accuracy with the epoch is the ratio between the output testing and the resulted output which is then divided by the total training result as what is given in the following pseudocode:

```
[m,n]= find (output == target);
accuracy= sum(m)/total_images*100
```

Examples of the output training illustrations and the output results can be seen in Figure 13.

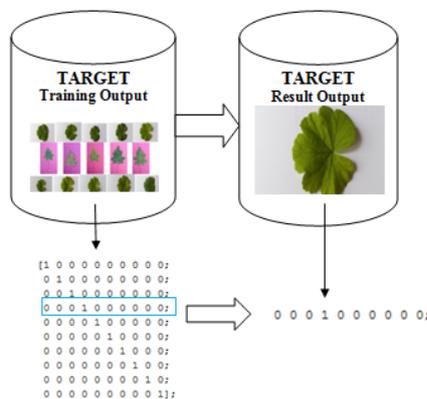


Figure 13. The illustration example of the output training and the output result of geranium leaf imagery

### 3. RESULTS AND ANALYSIS

#### 3.1. Experimental Image

Segmentation done to the plant class is based on the leaf feature texture as what can be seen in table 1. As can be seen in Table 1, the overall images have successfully been segmented well so that the edges of the leaf images are well segmented as what can be seen in Figure 13. Classification is done with reference to the species of each plant type as what can be seen in Table 3. As can be seen in Table 2, there is an error in the result of classification. The original image of a Croton is known as the Lychee type leaf image. This is due to the lack of species in the plant class dataset for a Croton so that the similarity of edges between a Croton and a Lychee results in such classification errors as what can be seen in Figure 14.

Table 1. The Leaf Image Segmentation

Plant Class	Binary Image	Segmentation Result
Geranium		
Ashanti Blood		
Bitter Orange		
Chocolate Tree		
EggPlant		
Ficus		
Croton		
Lychee		
Papaya		
Sweet Potato		

Table 2. Error in the Result of Classification

Plant Class	Original Image	Backpropagation Neural Network
	1	1
	2	2
	3	3
	4	2
	5	5
	6	6
	7	8
	8	8
	9	9
	10	10

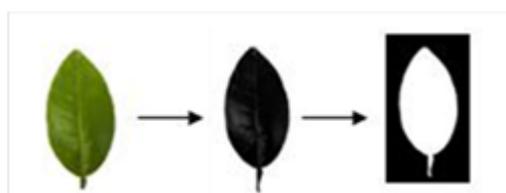


Figure 13. The image segmentation process of a bitter orange leaf

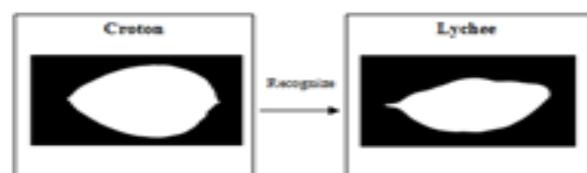


Figure 14. The example of the error in classifying a croton leaf

In Table 2, there is also an error in identifying the type of Chocolate Tree leaf that is identified as Ashanti Blood leaf. This is due to the similar value of the shape roundness of both species as what can be seen in Figure 15.

Table 3. The Number of Plant Classifications

Classification Number	Plant Classification
1	Geranium
2	Ashanti Blood
3	Bitter Orange
4	Chocolate Tree
5	EggPlant
6	Ficus
7	Croton
8	Lychee
9	Papaya
10	Sweet Potato

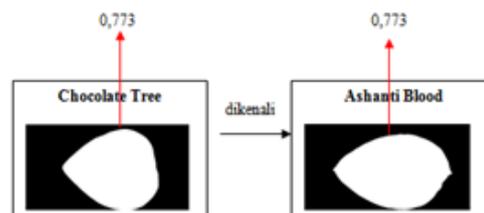


Figure 15. The example of the error in classifying a chocolate tree leaf

From the overall results of the classification training trials on 32 plant classes with 1605 plant leaf images there were classification errors in identifying 48 species out of 1557 species which were successfully identified with the resulted accuracy of 97% calculated by :

$$Accuracy = \frac{Perimeter}{\sum D} \times 100\% \quad (7)$$

Where P is a comparison between plant class species that are correctly classified by the total number of species. The researcher adjusted the connection weight during the training process of a number of datasets taken from the result of the extraction of the leaf feature characteristics to minimize the error value. In this study, the chain rule was implemented to calculate the influence of each weight on the error function in order to minimize the leaf image identification error.

#### 4. CONCLUSION

The segmentation of 1605 leaf images has been successfully done using the morphological operations. The extraction of the texture characteristics has also been successfully done based on the area value, the perimeter and such additional features of the leaf images as the roundness and the slimness of the leaf shape. The training of the extraction of leaf image characteristics was successfully performed using the backpropagation neural network. Based on the overall results of the classification testing trial on 32 classes of plants with 1605 plant leaf images there was a classification error of 48 species out of 1557 species which were successfully identified, resulting in an accuracy of 97%.

#### ACKNOWLEDGEMENTS

We are indebted to the experts who have contributed towards development of the template. The authors would like to acknowledge to Gunadarma University.

#### REFERENCES

- [1] Persoon and Weerd, "Biodiversity and Natural Resource Management in Insular Southeast Asia", *Island Studies Journal*, Vol. 1, No. 1, pp. 81-108, 2006.
- [2] Shankar, Ahmad, Pasrichaa and Sastry, "Bioreduction of Chloroaurate Ions by Geranium Leaves and Its Endophytic Fungus Yields Gold Nanoparticles of Different Shapes", *Journal of Materials Chemistry*, Issue 7, 2003.
- [3] Melly Br Bangun, Yeni Herdiyeni, Elis Nina Herliyana, "Morphological Feature Extraction of Jabon's Leaf Seedling Pathogen using Microscopic Image", *TELKOMNIKA (Telecommunication, Computing, Electronic and Control)*, Vol.14, No.1, Pp. 254-261, 2016.
- [4] Lembaga Ilmu Pengetahuan Indonesia, "Bioresources for Green Economy Development (in Bahasa)", Jakarta : LIPI, 2013.
- [5] Yan Qing, Liang Dong, Zhang Jingjing, "Research of Plant-Leaves Classification Algorithm based on Supervised LLE", *TELKOMNIKA (Telecommunication, Computing, Electronic and Control)*, Vol. 11, No. 6, pp. 3265-3270, 2013.

- 
- [6] Kadir, Nugroho, Susanto and Santosa, "Performance Improvement of Leaf Identification System Using Principal Component Analysis", *International Journal of Advanced Science and Technology*, Vol.44, 2012.
- [7] Kadir, Nugroho, Susanto and Santosa, "Leaf Classification Using Shape, Color, and Texture Features", *International Journal of Computer Trends and Technology*, 2011.
- [8] Resty Wulanningrum, Bagus Fadzerie Robby, "Learning Vector Quantization Image for Identification Adenium", *Indonesian Journal of Electrical Engineering and Computer Science (IJECS)*, Vol. 4, No. 2, pp. 383-389, 2016.
- [9] Chuan-Min Zhai, Ji-Xiang Du, "Applying extreme learning machine to plant species identification", *International Conference on Information and Automation (ICIA)*, 2008.
- [10] Leaf Dataset, <https://archive.ics.uci.edu/ml/datasets/leaf>, Acces Date : April 5, 2017
- [11] Susetianingtias D.T, Madenda, Rahayu D.A, Rodiah, "Retinal Microaneurysm Detection using Maximally Stable External Region Algorithm", *International Journal on Advanced Science Engineering Information Technology*, Vol. 6, No.5, ISSN : 2088-5334, 2016.
- [12] Gonzalez and Woods, "Digital Image Processing", Third Edition, Pearson Prentice Hall. ISBN 0-13-168728-x, 2008
- [13] Saveliev Peter, "Measuring objects, Computer Vision & Math contains: mathematics courses, covers: image analysis and data analysis, provides: image analysis software article", Available From : <http://inperc.com/>, 2011.
- [14] Abrham Debasu Mengistu, Seffi Gebeyehu Mengistu, Dagnachew Melesew, "An Automatic Coffee Plant Diseases Identification Using Hybrid Approaches of Image Processing and Decision Tree", *Indonesian Journal of Electrical Engineering and Computer Science (IJECS)*, Vol. 9, No. 3, pp. 806-811, 2018.