

## A Computer Aided System for Tropical Leaf Medicinal Plant Identification

Yeni Herdiyeni<sup>1#</sup>, Elvira Nurfadhilah<sup>1</sup>, Ervizal A.M. Zuhud<sup>2</sup>, Ellyn K. Damayanti<sup>2</sup>, Kohei Arai<sup>3</sup>, Hiroshi Okumura<sup>3</sup>

<sup>1</sup>*Department of Computer Science, Faculty of Mathematics and Natural Sciences  
Bogor Agricultural University, West Java, Indonesia  
E-mail: # yeni.herdiyeni@ipb.ac.id*

<sup>2</sup>*Department of Forest Resources Conservation and Ecotourism, Faculty of Forestry  
Bogor Agricultural University, West Java, Indonesia*

<sup>3</sup>*Graduate School of Science and Engineering,  
Saga University, Saga City, Japan*

**Abstract**— The objective of this paper is to develop a computer aided system for leaf medicinal plant identification using Probabilistic Neural Network. In Indonesia only 20-22% of medicinal plants have been cultivated. Generally, identification process of medicinal plants has been done manually by a herbarium taxonomist using guidebook of taxonomy/dendrology. This system is designed to help taxonomist to identify leaf medicinal plant automatically using a computer-aided system. This system uses three features of leaf to identify the medicinal plant, i.e., morphology, shape, and texture. Leaf is used in this system for identification because easily to find. To classify medicinal plant we used Probabilistic Neural Network. The features will be combined using Product Decision Rule (PDR). The system was tested on 30 species medicinal plant from Garden of Biopharmaca Research Center and Greenhouse Center of Ex-situ Conservation of Medicinal Indonesian Tropical Forest Plants, Faculty of Forestry, Bogor Agriculture University, Indonesia. Experiment results showed that the accuracy of medicinal plant identification using combination of leaf features increase until 74,67%. The comparative analysis of leaf features has been performed statistically. It showed that shape is a dominant features for plant identification. This system is very promising to help people identify medicinal plant automatically and for conservation and utilization of medicinal plants.

**Keywords**— leaf identification; leaf morphology; leaf shape; product decision rule; medicinal plants; probabilistic neural network.

### I. INTRODUCTION

Indonesia is a country with a mega biodiversity. For plants diversity, Indonesia has more than 38,000 plants species [1]. In 2001 the Laboratory of Plant Conservation, Faculty of Forestry Bogor Agricultural University (IPB) has recorded 2,039 tropical medicinal plant species from Indonesia forest ecosystems [2]. Unfortunately only 20-22% of medicinal plants are cultivated by people. One of the conservation and utilization of medicinal plants using the technology is developing a medicinal plants identification system automatically [3].

The plant identification is not easy and required an appropriate background with significant experience. Also number of researcher who can identify medicinal plants is very limited. The researcher must bring a book dictionary of

plant during plant identification in the field. This caused by utilization of medicinal plants by community is very low. Therefore, we need a computer-based automatically system as a tool to help people identify these various types of the medicinal plants.

There were some researches on plant identification. Fourier descriptor was used to extract leaf shape of Dicotyledonous leaf class [4]. The accuracy of system is 31.75%. Basic characteristic of leaf morphological and its derivatives have been analyzed by [5]. The accuracy of system is 27.22%. The Local Binary Pattern (LBP) is used to extract texture of house plants [6]. The accuracy of system is 73.33%.

In this research, we develop a computer aided system for Indonesian leaf medicinal plant identification using three

features, i.e., morphological, texture, and shape and also use Probabilistic Neural Network (PNN) for leaf identification.

## II. FEATURE EXTRACTION

### A. Morphological Feature Extraction

Morphological feature consists of two features, basic and derivate. The basic feature used in this research was diameter, area and perimeter/leaf circumference. Three basic features can be combined to get eight derivate features like smooth factor, shape factor, ratio perimeter and diameter, also five features of leaf vein [4].

### B. Texture Feature Extraction

Local Binary pattern was originally designed for texture description. The operator assigns a label to every pixel of an image using the 3x3-neighbourhood of each pixel with the center pixel value and considering the results as binary number. The histogram of the label is used as texture descriptor. Fig. 1 show illustration of the basic LBP operator.

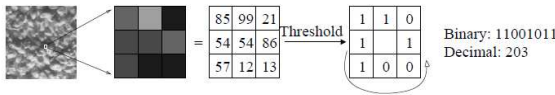


Fig. 1 Illustration of the basic LBP Operator

To be able to deal with textures at different scales, the LBP operator was later extended to use different sizes. Defining the local neighbourhood as a set of sampling points evenly spaced on circle centered at the pixel to be labelled allows any radius and number of sampling points. In this research the notation (P,R) will be used for pixel neighbourhoods which means P sampling points on a circle of radius R. Fig. 2 show an example of circular neighbourhoods.

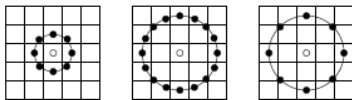


Fig. 2 Circular neighbourhoods with (8,1), (16, 2) and (8,2).

Generally, texture can be characterized by a spatial structure (e.g. a pattern such as LBP) and the contrast (e.g. Rotation Invariant Variance Measure (VAR), the variance of local image texture). Spatial structures vary with rotation while contrast does not. Ojala et.al [7] proposed using the joint histogram of the two complementary features, namely LBP/VAR, for rotation invariant texture classification. The drawback of this approach is that the value of VAR is continuous so that a quantization step is needed to calculate the histogram. LBPV is a simplified but efficient joint LBP and contrast distribution method. In this research we used LBPV with different operator.

### C. Shape Feature Extraction

To detect edge of leaf, this research used image thresholding and canny edge detector. Boundary of an object is detected using 4-connected neighbouring. Point by point

of boundary was traversed and represented as complex number,  $x + jy$ . Sequence of  $N$  complex number  $z(n) = x_n + iy_n$  was transformed using Fourier Transform, resulting  $Z(n)$ .  $Z(n)$  are called as Fourier descriptors.

Shape feature of *Fourier descriptors* result is expected to be invariant with the rotation, dilatation, and translation. For that purpose, some adjustments are needed on *Fourier descriptors* as result of the Fourier transformation. Translation is only affected to the first *Fourier descriptors* ( $Z(0)$ ), so the part that was took for the shape description is only the  $Z(n)$  line with  $1 \leq n \leq N$ . Rotation effect did not affected *Fourier descriptors* line. Dilatation effect still affected with same quantity for every *Fourier descriptors*. To remove scaling factor we made division of each of *Fourier descriptors* with one of family at  $Z(n)$  line, for example  $Z(1)$ .

### D. Probabilistic Neural Network (PNN)

PNN is an Artificial Neural Network (ANN) that uses radial basis function (RBF). RBF is a function that is shaped like a bell that scaling nonlinear variable [8]. PNN is composed of only three layers: the input layer, the pattern layer and the summation layers. The main advantage of using the architecture of PNN training data is easy and very fast. Weight are not "trained" but assigned. Existing weight will never be alternated but only new vectors are inserted into weight matrices when training. So it can be used in real time. In this research classifier combination is used to increase efficiency of decision making. Classifier combination used a combination of two or more value the individual characteristics of decision rules. Results of classification of each feature are produce the prior probability and posterior probability. Based on these probabilities, classifier combination techniques that can be used is product decision rule (PDR), sum decision rule (SDR), the maximum decision rule (MDR) and the majority vote rule (MVR) [9].

### E. Evaluation

Typically, performance system is mainly evaluated by accuracy of classification.

$$\text{accuracy} = \frac{\sum \text{testing data correctly classified}}{\sum \text{testing data}} \times 100\%$$

## III. EXPERIMENTAL RESULT

### A. Experiment Data

The entire document should be in Times New Roman or Times font. Type 3 fonts must not be used. Other font types may be used if needed for special purposes. Recommended font sizes are shown in Table 1.

The system was tested on 30 species medicinal plant from BioFarmaka garden and at Ex-situ Central Conservation of Tropical Forest medicinal plants Indonesia, Forestry Faculty, Bogor Agricultural University (IPB), West Java, Indonesia. The number of leaves is 1,440 images. Each of class consist of 48 images with resolution is 270 x 240 pixels. Species that we use in this research are Pandan Wangi / *Pandanusamaryllifolius Roxb* (class 1), Jarak Pagar/

*Jatrophacurcas Linn* (class 2), Dandang Gendis / *Clinacanthusnutans Lindau* (class 3), Lavende r/ *Lavendulaaffinialis Chaix* (class 4), Akar Kuning/ *Arcangelisiaflava L* (class 5), Daruju/ *Acanthus ilicifolius L* (class 6), Pegagan/ *Centellaasiatica, (Linn) Urban.* (class 7), Andong/ *Centellaasiatica,(Linn) Urban.* (class 8), kemangi/ *Ocimbasilicum* (class 9), Iler/ *Coleus scutellarioides, Linn, Benth*(class 10), jeruknipis/ *Citrus aurantifolia, Swingle.* (class 11), Bidani/ *QuisqualisIndica L.* (class 12), GadungCina/ *Smilax china* (class 13), Tabat Barito/ *Ficusdeloidea L.* (class 14), NandangGendisKuning (class 15), BungaTelang/ *Clitoriaternatea L.* (class 16), Mangkokan/ *NothopanaxscutellariumMerr.* (class 17), SomJawa/ *Talinumpaniculatum (jacq.) Gaertn.* (class 18), Pungpulutan/ *Urenalobata L.* (class 19), SosorBebek/ *Kalanchoepinnata(Lam.)Pers* (class 20), Nanas Kerang/ *Rhoeo discolor (L.Her.) Hance*(class 21), Seligi/ *PhyllanthusbuxifoliusMuell* (class 22), RemakDaging/ *Excecaria bicolor Hassk* (class 23), Kumis-Kucing/ *Orthosiphonaristatus (BI) Miq* (class 24), Kemuning/ *Murrayapaniculata [L.. ] Jack.* (class 25), CincauHitam/ *Mesonapalustris*(class 26), SambangDarah/ *ExcoceriacochinensisLour.* (class 27), Landik/ *BarlerialupulinaLindl.* (class 28), JambuBiji/ *Psidiumguajava L.* (class 29), and Handeuleum/ *Gryptophyllumpictum (L.) Griffith* (class 30).

### B. Classification of Morphological, Texture, Shape and Features Combination Using PNN

In this research, we conducted four experiments for plant identification, i.e., based on morphological feature classification, texture feature classification, shape features classification and combination of three features. In training phase, we used 38 pieces of leaves for each class (species) and 10 pieces of leaves for each class in testing phase. The comparative analysis of leaf features has been performed statistically.

#### Morphological Feature Classification

The experimental result showed that average accuracy of leaf morphological features only has the average accuracy is 20%. Some species get low accuracy. Fig 3 shows accuracy of leaf identification for each class using morphological feature.

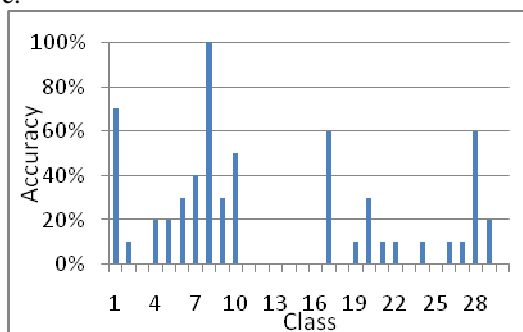


Fig. 3. Accuracy of plant identification based on leaf morphological feature.

Fig. 3 showed that morphological feature did not work for some species. According to observation, it is caused by morphological variety of species. Figure 4 showed sample of

species which has various morphological feature. Fig. 4 showed that on class 3 (DandangGendis) train data or test data images had various sizes with different angle, so that morphological feature especially basic leaf features can not be used for plant identification. For derivate leaf characteristic, it's quite good in representing the characteristic for leaf with the different size but on the same shoot angle.

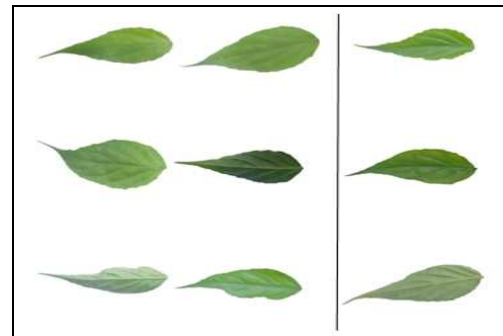


Fig. 4. Samples of image class 3 (DandangGendis) Texture

The experimental result showed that average accuracy of plant identification using LBPV (8,1) is 53%. Figure 5 showed accuracy of leaf identification for each class using LBPV. Some species can be identified correctly, for example class 9 and 12. But some species still has low accuracy caused by it has different illumination. Fig. 6 show sample of image (GadungCina) which has different illumination.

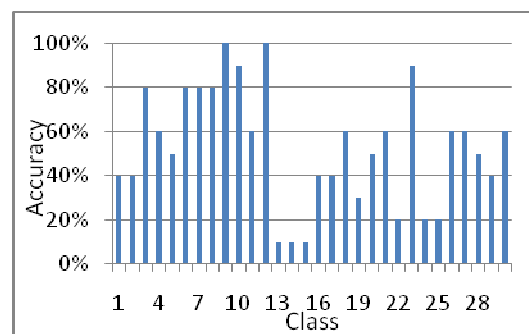


Fig. 5. Accuracy of plant identification based on LBPV (8,1).

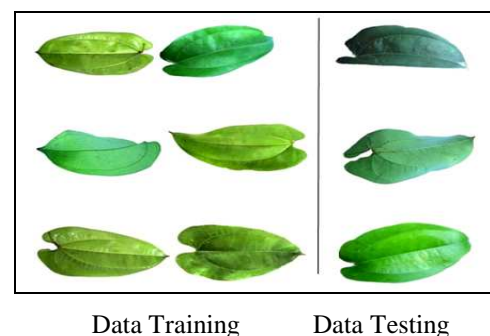


Figure 6. Samples of image class 13 (GadungCina).

#### Shape

The experimental result show that average accuracy of plant identification using shape feature is 64%. In this experiment we used Fourier Descriptor. Fig. 7 showed the

average accuracy for each class using shape features. The number of species which can be identified using shape feature more than species using texture or morphological texture.

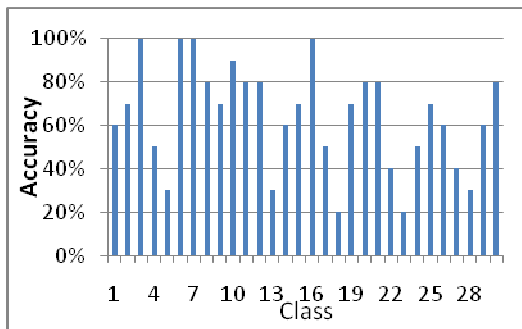


Fig. 7. Accuracy of plant identification based on Fourier Descriptor.

Fig. 8 showed sample of leaves which has low accuracy. The leaves have a various in shape so it is difficult to identify correctly.

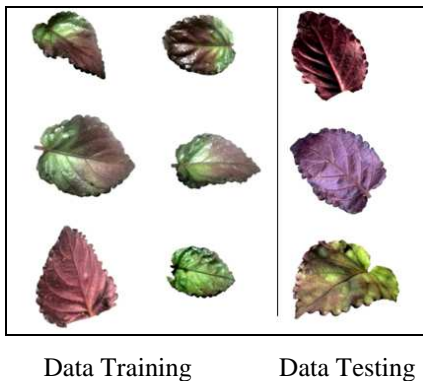


Fig. 8 Sample of image class 23 (RemakDaging).

### C. Feature Combination

The experimental result show that average accuracy of plant identification using feature combination increased up to 74.67% as shown in Fig. 9. The feature combination has better accuracy compare to single fature. In this experiment we used feature combination of morphology, texture and shape. In this research we used product decision rule (PDR) to combine these features. Fig. 11 showed the average accuracy for each class using feature combination.

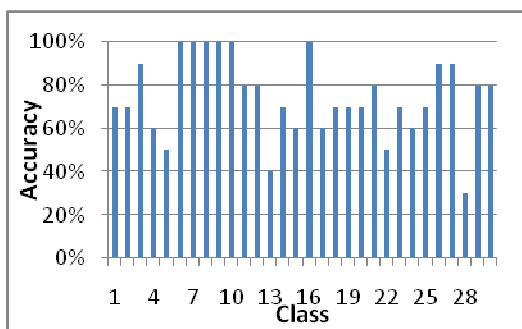
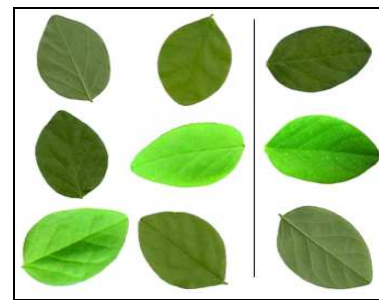


Fig. 9 Accuracy of plant identification based on feature combination.

Plant identification using feature combination can increase the accuracy of identification. It is showed that

different features can be used for plant identification. Fig. 10 show samples of image from class 16 (BungaTelang) which can be identified correctly using texture or morhology.



Data Training Data Testing

Fig. 10 Sample of image from class 16 (BungaTelang).

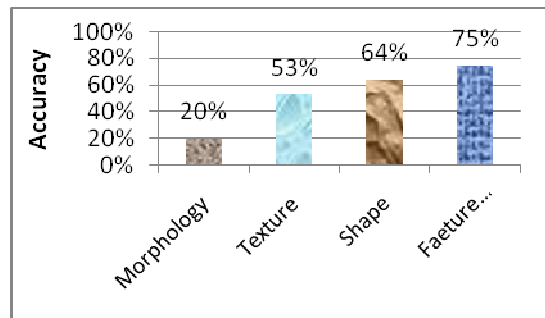


Fig. 11 Comparison of average accuracy using different features.

In this research we also performed statistically comparative analysis of leaf feature combination. It showed that shape is a dominant features in plant identification compare to morphology and texture as shown in Fig. 12.

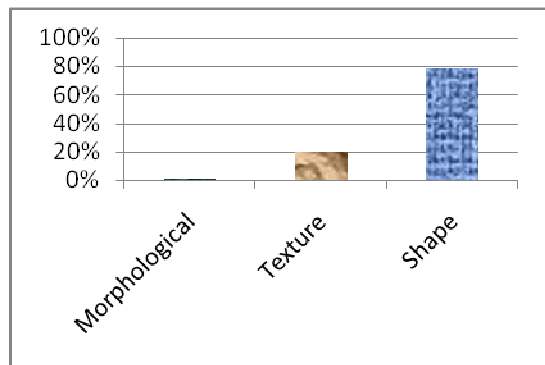


Fig. 12 Comparative analysis of leaf features.

### IV. CONCLUSIONS

A Computer aided system for tropical leaf medicinal plant identification has been successfully implemented. Plant identification is done based on leaf features combination (morphology, shape and texture). The classification result using the PNN method with the PDR classifier combination is quite good to increase the medicine plant identification accuracy up to 74.67%. The comparative feature analysis show that shape feature is most dominant than morphology and texture. This system is very promising to help people

identify medicinal plant automatically. Also this system can be used for conservation and utilization of medicinal plants.

#### ACKNOWLEDGMENT

The authors would like thank to Department of Forest Resources Conservation and Ecotourism, Faculty of Forestry, Bogor Agricultural University and Department of Information Science, Saga University, Japan for research collaboration. Their support is gratefully acknowledged

#### REFERENCES

- [1] Bappenas. 2003. Indonesia Biodiversity Strategy and Action Plan 2003-2020. Jakarta: Bappenas.
- [2] Zuhud, E.A.M. 2009. Potensi Hutan Tropika Indonesia sebagai Penyangga Bahan Obat Alam untuk Kesehatan Bangsa. Jurnal Bahan Alam Indonesia. Vol. VI No. 6, Januari 2009
- [3] Masyhud. 2010. Lokakarya Nasional Tumbuhan Obat Indonesia 2010. Jakarta : Pusat Penelitian dan Pengembangan Hutan Tanaman.
- [4] Herdiyeni, Y., and Ramadhani, M 2010. Shape and Vein Extraction on Leaf Image Using Fourier and B-Spline Model. Asian Federation for Information Technology in Agriculture, Bogor, Indonesia.
- [5] Annisa. 2009. Morphology and Texture Extraction for Leaf Retrieval. Thesis. Department of Computer Science. Bogor Agricultural University.
- [6] Kulsum, Lies U. 2010. House Plant Identification Automatically using Local Binary Patterns Descriptor and Probabilistic Neural Network. Thesis. Department of Computer Science. Bogor Agricultural University.
- [7] Ojala T., et al. 2002. Multiresolution Gray-Scale and Rotation Invariant Texture Classification with Local Binary Pattern. IEEE Transactions on PAMI. Vol. 24, No. 7, pp. 2037-2041.
- [8] Wu S. G., et al. 2007. A Leaf Recognition Algorithm for Plant Classification Using Probabilistic Neural Network. China : Chinese Academy of Science.
- [9] Kittler, J., et al. 1998. On Combining Classifiers. IEEE Transactions On Pattern Analysis And Mechine Intelligence. Vol 20 no.3 page : 226-239.