Journal Of Engineering And Technology Innovation (JETI)
June 2022 | Vol. 1 | No. 2
E-ISSN : 2828-1209
E-ISSN : 2828-1462

Medicinal Plant Processing Viewing Leaf Morphological Feature Extraction With Artificial Neural Tissue

David Saro *[1]
[1]Universitas Ibnu Sina
e-mail: *[1]dsaro89s@gmail.com,

Abstract

Indonesia has long known and used plants that are efficacious as medicine. Of the many medicinal plants in the world, 80% of medicinal plants grow in tropical forests in Indonesia. About 28,000 plant species grow and 1,000 species of them have been used as medicinal plants. The large number of medicinal plant species and the high degree of similarity can cause errors in the process of identifying medicinal plant species. So it takes the help of a computer to identify the types of medicinal plants. The purpose of this study was to identify types of medicinal plants using backpropagation neural networks based on leaf morphological feature extraction. The results show that changes in the learning rate value affect the results of the identification of medicinal plant species based on leaf morphological features. The results of the calculation of the average recognition rate value of 90% for training data and 75.56% for testing data occur when the learning rate is 0.01. The best learning rate value for the identification of medicinal plant species is 0.01 with an average number of epochs of 11.67 and MSE of 0.13. This shows that leaf morphological feature extraction methods and backpropagation neural network algorithms are very well used to identify medicinal plant species.

Keywords— *Feature Extraction, Backpropagation Neural Networks, Leaf Morphology, Medicinal Plants*

INTRODUCTION

Indonesia has long known and used plants that are efficacious as medicine. Medicinal plants are used as an effort to overcome health problems. Knowledge of medicinal plants is based on experience and skills that have been passed down from generation to generation (Sari, 2006). Tropical forest species in Indonesia are estimated at 143 million ha. Tropical forests are where 80% of medicinal plants grow in the world, where 28,000 plant species grow and 1,000 species of them have been used as medicinal plants.

The large number of medicinal plant species and the high degree of similarity can cause errors in the process of identifying medicinal plant species. Misidentification of medicinal plant species can be fatal for those who consume it, and can even result in death. So it takes the help of a computer to identify the types of medicinal plants. Extraction of leaf image features based on morphological features has been used as a marker to identify the type of disease in Jabon plants with the results showing that morphological features can represent the characteristics of leaf diseases in Jabon plants.

In addition, leaf morphological features have also been used to identify plant species using the k-Nearest Neighbors classification algorithm with an accuracy of 92% (Saragih & Wahyuni, 2018). Based on this research, it shows that morphological features can be used to characterize a leaf. Research on the application of artificial neural network methods has been carried out to classify cardiotocography with a high accuracy value of 99.15%.

In addition, the artificial neural network method has also been applied to classify soil with accuracy results for 2 classes of 81.89%, 3 classes of 83.40%, and 7 classes of 63.40% (Mukminin & Riana, 2017). From this research, it shows that the artificial neural network method is quite good in classifying. In this study, the classification of medicinal plants uses an artificial neural network classification algorithm using an image dataset of medicinal plant leaves.

RESEARCH METHODS

The leaf images of 75 medicinal plants were divided into 60 training data and 15 testing data. Each type of medicinal plant has 12 training data and 3 testing data as shown in Table 1. The image background is white, so it will not affect the feature extraction results because feature extraction is carried out based on leaf morphological features, not leaf color.

Data Preprocess

After the leaf image data is obtained, the next step is to preprocess the leaf image. There are 2 stages to preprocess the obtained leaf image data. The first process is leaf image cutting (cropping) and the second is leaf image segmentation. Image cropping is done to obtain a full leaf image as shown in Figure 1 (a). While image segmentation aims to separate the object (foreground) from the background. The result of image segmentation is a binary image where the desired object (foreground) is white, while the background is the larger the rectangularity value of an object, the object is in the form of a box. which you want to remove is black as shown in Figure 1 (b). Segmentation is a very important process in object recognition. The segmentation method used is thresholding. The feature extraction process is carried out after the segmentation process is complete.

Figure 1. Preprocessing

Feature Extraction For extraction of morphological features, there are several features that can be calculated, such as area and perimeter. Based on the calculation of the area, perimeter, major axis, and minor axis, other morphological features can also be calculated. Following are some of the formulas used to extract morphological features:

1. Eccentricity (Ecc)
   It is the ratio of the distance between the foci of the ellipse to the length of the major axis of an object. Eccentricity is worth between 0-1. Eccentricity is a technique to describe an object with an ellipse shape.

$$Ecc = \frac{\sqrt{(major\ axis^2 - minor\ axis^2)}}{major\ axis} \quad (1)$$

2. Solidity (Sole)
   Measuring the density of an object, solidity is the ratio of the object's area to the full convex area of the object.

$$Sol = \frac{area}{convex\_area} \quad (2)$$

3. Rectangularity (Rect)
   A technique to describe the similarity of an object's shape to a box shape. The greater the rectangularity value of an object, the object is in the form of a box.

$$Rect = \frac{area}{major\ axis\ x\ minor\ axis} \quad (3)$$

4. Metric (Met) Is the ratio between the area of the object with the area of a circle using the same perimeter. Metric is a technique to describe the compactness of an object.

$$Met = \frac{4\ x\ \pi\ x\ area}{perimeter^2} \quad (4)$$

5. Extension (Ext)
   The proportion of pixels in the bounding box that are in the a

$$Ext = \frac{area}{area\ of\ the\ bounding\ box} \quad (5)$$

6. Elongation (Elong)
   Measure the slenderness of an object. If the value obtained is close to 1 then the object has an elongated shape.

$$Elong = 1 - \frac{minor\ axis}{major\ axis} \quad (6)$$

Data Sharing

The distribution of data from leaf morphological feature extraction using the k-fold cross validation method. The number of k used is 3. Given the small amount of data used for training, the selection of 3-fold cross validation is sufficient to make data variations, so that all data is used, both for test data and training data. The entire feature extraction data is divided into 3 subsets, namely D1, D2, D3. Each subset has the same size. In the first process D2, D3 becomes training data and D1 becomes test data, in the second process D1, D3, becomes training data and D2 becomes test data, and so on (Manik & Saragih, 2017).

Artificial Neural

Network This research uses backpropagation neural network algorithm. This method works through an iterative process using 172 JOURNAL INFORMATIKA Vol. 5 No.2, September 2018: 169 – 174 set of sample data (training data), comparing the predicted value of the network with each example. In each process, the relation weights in the network are modified to minimize the Mean Square Error (MSE) value between the predicted value of the network and the real value.

Modification of the artificial neural network relation is carried out in a backward direction, from the output layer to the first layer of the hidden layer, so this method is called backpropagation (Novita, 2016). The backpropagation algorithm uses error output to change the values of the weights in the backward direction. To get this error, the forward propagation step must be done first. At the time of forward propagation, the neurons are activated using the sigmoid activation function (Dalimunthe, 2016).

The use of the learning rate parameter for each case is different, as was done in the case of the classification of diabetes which resulted in the best learning rates of 0.01, 0.02, and 0.03 (Pandu Cynthia & Ismanto, 2017). This research uses learning rate parameters of 0.01, 0.02, and 0.03.

HASIL DAN PEMBAHASAN

Feature Extraction

Samples extracted from leaf morphological features for medicinal plants can be seen in Table 2 and Table 3.

Table 1 Feature Extraction Sample

Leaf	Feature Extraction		
	Ecc	*Rect*	*Sol*
Binahong	0,5324	0,7546	0,9669
Jambu	0,9157	0,7794	0,9852
Keji Beling	0,7858	0,7822	0,9757
Sirih	0,6934	0,7668	0,9682
Som Jawa	0,8595	0,7608	0,9797

Table 2 Advanced Feature Extraction Sample

Leaf	Feature Extraction		
	Met	*Ext*	*Elong*
Binahong	0,5839	0,7094	0,1535
Jambu	0,5431	0,7435	0,598
Keji Beling	0,4629	0,6938	0,3815
Sirih	0,672	0,6652	0,2795
Som Jawa	0,5691	0,6805	0,4889

Artificial Neural Network Implementation

This study uses the binary sigmoid activation function (logsig) in the hidden layer and output layer. While the training function uses the Levenberg Marquardt Backpropagation method. The number of hidden neurons used is 5, the maximum number of epochs is 100, and the learning rate is 0.01, 0.02, and 0.03. The results of the study will see the effect of changes in learning rate and fold on Mean Squared Error (MSE) and Recognition. Rates (%). The results of training data processing can be seen in Table 4.

Table 3. Results of Training Data Processing

Fold	Learning Rate	Results		
		Amount epoch	MSE	*Recognition Rate* (%)
1	0,01	12	0,12795	78,33
	0,02	13	0,64714	91,67
	0,03	9	0,77411	83,33
2	0,01	11	0,22635	96,67
	0,02	15	0,43226	91,67
	0,03	16	0,78704	93,33
3	0,01	12	0,030178	95
	0,02	10	0,55377	88,33
	0,03	13	0,57525	61,67

Table 4 shows that changes in the value of learning rate and the number of folds affect the results of the calculation of the recognition rate. The highest recognition rate calculation occurs at fold 2, the learning rate is 0.01, and the MSE is 0.22635, which is 96.67%. To avoid overfitting, the average recognition rate is calculated for each learning rate. The results of the calculation of the average recognition rate for training data can be seen in Table 5.

Table 4. Average Calculation Results

Learning Rate	*Epoch*	MSE	*Recognition Rate* (%)
0,01	11,67	0,13	90,00
0,02	12,67	0,54	90,56
0,03	12,67	0,71	79,44

Table 5 shows that the highest average recognition rate occurs when the learning rate is 0.02, which is 90.56%. However, based on the consideration of the average number of epochs and MSE, the best recognition rate value of 90.00% is taken when the learning rate is 0.01. To prove the 0.01 learning rate is the best, then a test is carried out using testing data. The results of testing data processing can be seen in Table 6.

Table 5 Results of Testing Data Processing

Fold	*Learning Rate*	*Recognition Rate (%)*
1	0,01	80
	0,02	80
	0,03	66,67
2	0,01	66,67
	0,02	73,33
	0,03	73,33
3	0,01	80
	0,02	46,67
	0,03	46,67

To adjust the calculation when using training data, the average recognition rate is calculated for each learning rate in the testing data. The results of the calculation of the average recognition rate for data testing can be seen in Table 7.

Table 6 Average Calculation Results

Learning Rate	Rata-rata *Recognition Rate* (%)
0,01	75,56
0,02	66,67
0,03	62,22

Based on the calculation of the average recognition rate in Table 7, it can be seen that the 0.01 learning rate is the best with the recognition rate calculation of 75.56%. This shows that the best learning rate for the identification of medicinal plant species is 0.01 with an average number of epochs of 11.67 and MSE of 0.13. The average identification error occurred in the leaves of Binahong, Betel, and Som Jawa.

CONCLUSION

Leaf morphological feature extraction methods and backpropagation neural network algorithms are very well used to identify medicinal plant species with the calculation results of an average recognition rate of 90.00% for training data and 75.56% for testing data. Changes in the learning rate value affect the results of the identification of medicinal plant species based on the morphological features of the leaves. The best learning rate value for the identification of medicinal plant species is 0.01 with an average number of epochs of 11.67 and MSE of 0.13. It is hoped that this research can provide information about the parameters that can be used in the backpropagation neural network algorithm. To increase the accuracy obtained, it is possible to try different parameters and increase the amount of image data on medicinal plant leaves.

REFERENCES

Dalimunthe, F. H. (2016). Perancangan aplikasi mengidentifikasi penyakit mata dengan menggunakan metode backpropagation. Jurnal Riset Komputer, 3(1), 7–11.

Manik, F. Y., Herdiyeni, Y., & Herliyana, E.N. (2016). Leaf Morphological Feature Extraction of Digital Image Anthocephalus Cadamba.TELKOMNIKA (Telecommunication Computing Electronics and Control), 14(2),630. https://doi.org/10.12928/telkomnika.v14i2.2675

Manik, F. Y., & Saragih, K. S. (2017). Klasifikasi Belimbing Menggunakan Naïve Bayes Berdasarkan Fitur Warna RGB. IJCCS (Indonesian Journal of Computing and Cybernetics Systems), 11(1),99.https://doi.org/10.22146/ijccs.17838

Mukminin, A., & Riana, D. (2017). Komparasi Algoritma C4 . 5 , Naïve Bayes Dan Neural Network Untuk Klasifikasi Tanah. Jurnal Informatika, 4(1), 21–31.

Novita, A. (2016). Prediksi Pergerakan Harga Saham Pada Bank Terbesar Di Indonesia Dengan Metode Backpropagation Neural Network. Jutisi, 05(01), 965–972.

Arifin, N. Y., Kom, S., Kom, M., Tyas, S. S., Sulistiani, H., Kom, M., ... & Kom, M. (2022). Analisa Perancangan Sistem Informasi. Cendikia Mulia Mandiri.

Pandu Cynthia, E., & Ismanto, E. (2017). Jaringan Syaraf Tiruan Algoritma Backpropagation dalam Memprediksi Ketersediaan Komoditi Pangan Provinsi Riau. Rabit, 2(2), 196–209. Retrieved from
http://jurnal.univrab.ac.id/index.php/rab it/article/view/152