Leaf Features Based Plant Classification Using Artificial Neural Network

ManishaAmlekar[1], Dr. Ramesh R Manza[2],
Dr. Pravin Yannawar[3] and Dr. Ashok T Gaikwad[4]
[1, 4] *Institute of Management Studies and IT, Vivekananda College Campus, Aurangabad*
[2,3] *Department of CSIT Dr. BAMU, Aurangabad.*

ABSTRACT

This paper presents the classification of plant leaf images with biometric features. Traditionally, the trained taxonomic perform this process by following various tasks. The taxonomic usually classify the plants based on flowering and associative phenomenon. It was found that this process was time consuming and difficult. The biometric features of plants leaf like venation make this classification easy. Leaf biometric feature are analyzed using computer based method like morphological feature analysis and artificial neural network based classifier. ANN model take input as the leaf venation morphological feature and classify them into four different species. The result of this classification based on leaf venation is achieved 96.53% accuracy in the training of the model for classification of leaves provide the 91% accuracy in testing to classify the leaf images.

Keyword - Artificial neural network, Canny edge detection, k-NN Classification, Leaf venation pattern, Morphological Features.

Introduction

Plants are an integral part of ecosystem. Due to deforesting many plant species are under the risk of annihilation. Plants are useful for human being and other living things. They are useful as foodstuff, as medicine and also in many industries. Identifying plants helps ensure the protection and survival of all natural life. Plant identification can be performed using many different techniques using the plant's leaves. Leaves are useful to classify plants since they are more readily available than the other biometric components like flowers which are available for a short period. Plant classification by using leaves requires different biometric features of leaf like color, shape, texture and venation. This identification manually is time consuming and expensive. Leaves can be classified based on color that include similarity between two images with the help of color histogram, but the color based classification is depend on season and effect of sunlight. Artificial neural network is a system closely modeled on the human brain. Artificial neural network contains the multiple layers of simple processing elements called neuron. Each neuron is linked to certain of its neighbors with coefficients of connectivity that represent the strengths of these connections. Learning is accomplished by adjusting these strengths to cause the overall network to output appropriate results. Artificial neural network is trained with six nodes of the multilayer perceptron learning method for input to the neural network. Each node take input as the leaf venation morphological feature and classify the leaf images into four different species. This paper is organized into four sections Introduction, Previous work, Methodology, Conclusion.

Previous work

There has been substantial work in recent years in the field for leaf biometric recognition. Initially it was approached by Petry. He classified weed species based on shape and structure of leaves to automatically. This morphological feature extraction technique has been used many times. (Stephen Gang Wu , 2007) used twelve morphological features (including vein features) and a neural network to achieve 90.3% classification accuracy. By using eleven morphological features Knight (2010) achieved a classification accuracy of 80%. This approach was implemented as a mobile application designed for field guides. S.Prasad (2011) has worked on leaf recognition using support vector machine with relative sub image based features. T. Beghin (2010) used a number of morphological features and a fuzzy surface selection technique to achieve 99% classification accuracy even with leaves that are deformed and oriented incorrectly. Madhusmita Swain (2012) performed plant classification by using the full colour of the leaf in conjunction with support vector machine achieved 95% accuracy. However, this has the disadvantage that dry leaves cannot be used as they will be a different colour from a leaf that has just been picked. Ehsanirad (2010) used a Gray-Level Co-occurrent Matrix (GLCM) and Principal Component Analysis (PCA) to achieve classification accuracies of 78.46% and 98.46% respectively. Using a Probabilistic Neural Network with, K. Singh (2010) achieved a 91% classification accuracy. J. S. Cope (2010) compared four different texture methods for classification Gabor filters, Fourier descriptors, Co-ocurrent matrices and Gabor Co-occurrences to achieve classification with accuracies of 50.78%, 82.42%, 69.14% and 85.16% respectively. Finally, some researchers have combined both morphological and texture-based techniques. T. Beghin (2010) worked on Contour signature method for shape classification and the Sobel operator for texture classification. The result was a classification rate of 81.1%, significantly better than either of the methods when used alone. The problem with a number of these techniques is that they require some manual intervention such as correctly orienting the image or identifying the end points of the leaf's main vein. S. R. Deokar (2013) has worked on leaf recognition by extracting 28 and 60 Feature point. These features extracted by vertical and horizontal splitting the leaf images. ANN is used to compare performance of leaf recognition.

Methodology

Images of the leaves can be acquired many ways such as taking photo snap or by scanning the leaf using scanner. The experiment carried out in this research work is based on the image database published by Intelligent Computing Laboratory, Chinese Academy of Sciences 2010-2012. This research work experiment was framed on 210 leaves samples of four different species. The morphological features are used in this experiment to find the leaf venation patterns.

Flow of the methodology:

1. Preparation of input data images
2. Find the leaf venation pattern
3. Capture the leaf venation features
4. Classify leaf venation features

Preparation of Input data images

Leaf image is collected from the ICL database. Sample leaf image is as shown in fig1(a).

a) b) c)

Figure-1: Leaf Images a) original b) gray scale c) venation pattern

As the color of the leaf can be affected by the sunlight and season, this leaf image is to be processed to be in uniform form. For this purpose, leaf image is converted to gray scale by using the formula:

$$Gr=0.21R+o.71G+o.o7B$$

where Gr is corresponding pixel gray value, R is red color component value, G is green color component value, B is blue color component value. The gray scale form of the image is shown in fig1 (b).

Find the leaf venation pattern

By using the canny edge detection method the venation pattern of the leaf is found. This method first perform the smoothing the image to remove the noise. Then veins are identified only with local maxima, by finding the gradients of large magnitude. It finds the image gradient to highlight veins with high spatial derivatives. It determines the potential edges by double thresholding. Vein detection is finalized by suppressing all veins that are not connected to strong veins. Gradient magnitude is calculated by using formula

$$|G|=\sqrt{Gx^2+Gy^2}$$

Approximate magnitude $|G|=|Gx|+|Gy|$

The direction of the edge is computed using the gradient in the x and y directions. For finding the vein direction it uses the formula *Theta = invtan (Gy / Gx)*

Canny edge detection algorithm is good for finding the venation pattern as it suppress the veins that are loosely connected.

Capture the leaf venation features

For the leaf vein patterns the following leaf venation morphological features are captured

Area - Area of leaf in the region of leaf.

Eccentricity - The eccentricity of the elliptical shape of the leaf is the ratio of the distance between the foci of the elliptical shape and its major axis length.

Perimeter - Perimeter is the length of the outside boundary of the leaf. That means it is the distance around the boundary of the region. This is computed by calculating the distance between each adjoining pair of pixels around the border of the leaf.

Major Length- Major Length is the length of the major axis of the elliptical shape that has the same normalized second central moments as the region.

Minor Length- Minor Length is the length of the minor axis of the elliptical shape that has the same normalized second central moments as the region.

Convex Area- This is an area of the convex leaf image, which specifies the number of pixels in convex image.

Solidity- Solidity is also known as the convex area ratio calculated using:

$$Solidity = Area/ConvexArea$$

Classify the leaf venation features

By two different methodologies are followed to classify the leaf venation features

4.1 By using the k-nearest neighbor classifier

4.2 By using back propagation artificial neural network.

K-nearest neighbor classifier

k- nearest neighbor classifier the training leaf venations feature samples are described by n-dimensional numeric attributes. The training leaf samples are stored in an n-dimensional space. When a test leaf sample is given, the k-nearest neighbor classifier searches the k training leaf samples which are near to the unknown sample. Closeness is usually defined in terms of Euclidean distance. The euclidean distance is between two points P(p1,p2, …. Pn) and Q(q1,q2,…. qn) given by

$$D(P,Q)=\sqrt{\sum_{i=1}^{n}(P_i - Q_i)^2}$$

This KNN model is trained with total 173 leaf image venation features with 83% percentage of classification and 46 samples are tested with 54% of classification. Results of the classification are as shown in the following table1. This model provides 85.54% of accuracy in training the model and 65.21% accuracy for testing.

Table 1: Result of the KNN classification

Result	Training mode	Testing mode
Number of Leaf Images	173	46
Correctly Classified Images	145	25
Percentage of Classification	83.81 %	54.35 %
Mean absolute error	0.1169	0.2186
Root mean squared error	0.2367	0.3815

Table 2: Result of the ANN classification

Result	Training mode	Testing mode
Number of Leaf Images	173	46
Correctly Classified Images	157	37
Percentage of Classification	90.75 %	80.43 %
Mean absolute error	0.0866	0.1266
Root mean squared error	0.1948	0.2796

Artificial Neural Network classifier:

The artificial neural network with multilayer architecture is followed for the classification. In back propagation neural network approach, functional signals flows forward and error signals propagates backward. The artificial neural network is multilayer perceptron model that includes input layer, hidden layer and output layer. On input layer the leaf venation pattern features are provided as input to the model. Hidden layer of the model performs sigmoid activation function. This function has the formula $f(x)=1/1+e^{(-x)}$. This generates positive value in between 0 and 1.Output of the hidden layer is provided to output layer for classification of the leaf images in four different species classes. Seven morphological features of leaf venation pattern for 173 leaves of four different plant species with their weights are input to the model. The model provides the features classified into four classes for four species as output. The hidden layer performs the sigmoid function, error signals if generated in the processing at the nodes they are propagated back to the back node of the model. Accordingly the node update the weight the model in the training mode classifies the leaf images with 96.53% of classification. The model in testing mode, classify the 46 different leaf images of four species with 91% of classification results of the model is as shown in table2.

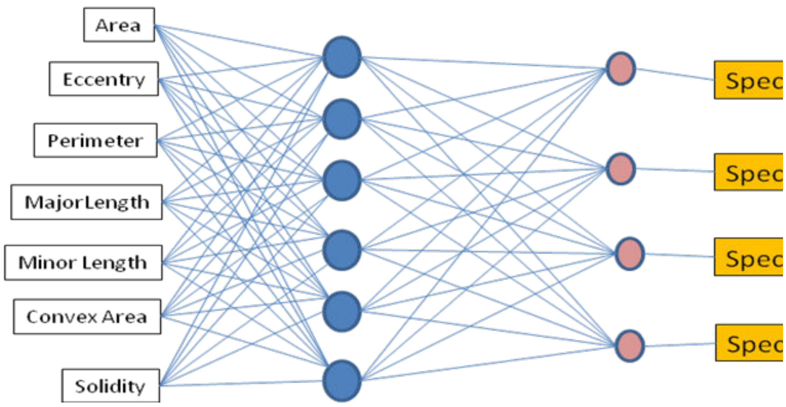

Figure-2: Artificial Neural network model

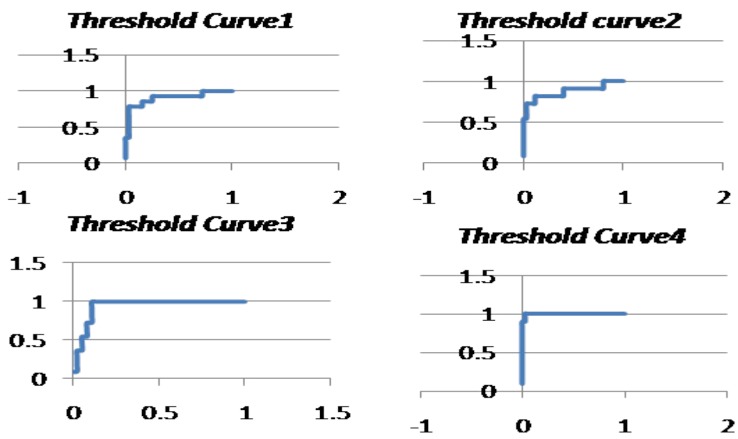

Figure-3: Scatter plot for ANN classification.

The multilayer perceptron learning model is designed as in Fig2 where model is trained with 10 nodes 500 epochs. This model uses one hidden layer that reduces the complexity of the model Fig 3 shows the scatter plot for the threshold curve for four classes of the species classification. Scatter plot shows the curve for the classes defined for four species by ANN model. Threshold curve towards the left top indicate the ratio of false positive rate vs. true positive rate of the classification. It is obtained here that the classification model provides the accuracy 96.53% in training and 91% accuracy for testing the samples.

Conclusion

The results are obtained based on total 221 leaf images. Leaf Images are processed by applying canny edge detection and morphological feature extraction methods. The results of the classification model are obtained by following multilayer perceptron based artificial neural network method.

It is obtained that plant classification based on leaf venation morphological features with KNN classification technique is providing 85.54% accuracy in training and 65.21% accuracy in testing mode. The ANN based classification model trained with **96.53%** accuracy provides the **91%** accuracy in testing to classify the leaf images with their morphological vein biometric features.

From the result it is clearly seen that ANN classification model has good classification performance than KNN classification model

Acknowledgement

We specially acknowledge to Intelligent Computing Laboratory, Chinese Academy of Sciences as they provided the database for this research work.

References

Amlekar Manisha, Manza R.R, Yannawar Pravin,(2013), Leaf classification based on leaf dimension biometric features of leaf shape using k-means classifier, *NCAC*, Jalgoan.

Amlekar Manisha, Manza R.R, Yannawar Pravin, Gaikwad B.P,(2013), Image data mining for classifying leaf dimension biometric features of leaf shape using KNN classification technique, *CMS*, Aurangabad.

Beghin T., Cope J. S., Remagnino P. and Barman S., (2010), Shape and texture based plant leaf Classification, *ACIVS*, 2, 345–353.

Cope J. S., Remagnino P., Barman S., and Wilkin P., (2010), Plant texture classification using gabor co-occurrences," in Proceedings of the *6th international conference on Advances in visual computing,*2, 669– 677.

Deokar S. R., Zope P. H. Suralkar S. R ,(2013), Leaf Recognition Using Feature Point Extraction and Artificial Neural Network, *International Journal of Engineering Research & Technology (IJERT)*, ISSN: 2278-0181, 2(1).

Ehsanirad, (2010), Plant classification based on leaf recognition, *International Journal of Computer Science and Information Security*, 8(4), 78–81.

Knight D., Painter J., and Potter M, (2010), Automatic plant leaf classification for a mobile field guide, *Stanford*, Tech. Rep..

Prasad S., Kudiri K. M., and Tripathi R. C, (2011), Relative sub-image based features for leaf recognition using support vector machine, in Proceedings of the *International Conference on Communication, Computing Security, ser. ICCCS*,343–346.

K. Singh, I. Gupta, and S. Gupta,(2010) ,Svm-bdt pnn and fourier moment technique or classification of leaf shape, *International Journal of Signal Processing, Image Processing and Pattern Recognition*, 67– 78.

Swain Madhusmita, Dash Sanjit Kumar, Dash Sweta and Mohapatra Ayeskanta,(2012), An approach for iris plant classification using neural network, *International Journal on Soft Computing (IJSC),*3,(1).

Madhusmita Swain, Sanjit Kumar Dash, Sweta Dash and Mohapatra Ayeskanta,(2005), An approach for iris plant classification using neural network, *International Journal on Soft Computing(IJSC)*,3(1).

Wu Stephen Gang, Forrest Sheng Bao, Xu Eric You, Wang Yu-Xuan, Chang Yi-Fan and

Xiang Qiao-Liang,(2007),Leaf Recognition Algorithm For Plant Classification Using Probabilistic Neural Network, *IEEE International Symposium On Signal Processing And Information Technology.*