Plant Classification Based on Leaf Features

Manisha Amlekar[1], Dr. Ramesh R Manza[2], Dr. Pravin Yannawar[3] and Dr. Ashok T Gaikwad[4]

[1,4] *Institute of Management Studies and IT, Aurangabad*

[2,3] *Department of CSIT Dr.Babasaheb Ambedkar Marathwada University,Auranagabad*

ABSTRACT

Plants are very much significant component of ecosystem. Plants can be used as foodstuff, in medicines and in many industries for manufacturing various products. Identifying those helps ensure the protection and survival of all natural life. Plant identification can be performed using many different techniques. As plant leaves are more readily available, it is efficient to identify and classify plants by using their leaves. Plant classification by using leaves requires different biometric features. This paper presents the review on various methods for plant classification based on leaf biometric features. For plant classification traditionally, the trained taxonomist and botanist had required to perform set of various tasks. The taxonomist usually classifies the plants based on flowering and associative phenomenon. It was found that this process was time consuming and difficult for following various tasks. The biometric features of plants leaf such as shape and venation make this classification easy. This review study may help the rural people for easily identifying in addition to classifying the plant based on the leaf features. This plant classification method include two basic tasks leaf biometric feature extraction and classification of plants based on these features.

Key words: Artificial Neural Network, Leaf Shape, Leaf Venation, Morphological Features, Probabilistic Neural Network

Introduction

Plants are important sources for human being and many other living as food stuff. They are also important for development in industries, for manufacturing of various home appliances, for using as ingredients in medicine. It is also important for balancing the ecosystem and environmental protection. There are 50 to 70 thousand species are known globally. [5] Many of the plant species are yet unknown and due to environmental imbalance many unknown species might be on the boundary of extermination [9]. India covers more than 45,000 species of flora, out of which several species are, not found anywhere else. Therefore it necessary for preserving these natural resources and hence, it is great requirement to correctly and quickly

identify plant species and classify them accordingly. The classification process of plant species identifies difference kinds of plants. Thus builds the system to classify the plant. Many researchers have tried to classify the plat species based on the various features of plants. Many of them have used leaf biometric features for plant classification as they are more readily available in almost all season. In the classification system plants are classified on the basis of color, shape venation and texture features. For identification of the plant species, individual plants are classified to a species based on its leaf physiological characteristics. Plant taxonomy methods identifies the plants by studying internal structure of plants called plant anatomy, or by identifying external functional study of various physical components. Plants can also be identified by cellular structure or molecular structure. This work, usually done by botanists, and trained taxonomist. Process of performing these tasks is very troublesome and time consuming not much efficient. This is very much important to improve this process by identifying plant species by using computerized techniques. Many researchers are started by using the methods that uses k-nearest neighbor (k-NN) classifier and adopted Artificial Neural Network (ANN) [2]. J.Reddy L. Manjunatha Rao, and M. Nachappa has presented study on multimodal biometric system. In this paper presented here the issues related to multi algorithm biometric systems. The performance of multi algorithm system is presented for improvement of multimodal biometric system. Various fusion levels and scenarios of multi algorithm systems are discussed [6]. Plant taxonomy theory says that leaves and flowers are more important components of plant biometry that are essentially used for classifying plants. Leaves are the components readily available in almost all season. Therefore are used in plant classification. Leaves are important part of plant and are distinct in shape and venation. A leaf is containing three basic pars, leaf lamina, stem and a small leaf like component at the base of the stem called stipules. Generally, flat or broad surface of the leaf called lamina or leaf blade is considered for plant classification techniques that use digital method for plant classification. The blade has a distinct shape and venation pattern. The leaf is used for classification as it is difficult to analyze flower shapes and structures as they have complicated and available only for limited duration. Shape is an important image feature perceived by the human to characterize an object [5] According to the plant taxonomy, leaf shape is important and effective for classifying the plants [7]. Many morphological differences exist in different kinds of leaves. Leaves are used for the classification of plants. The process of observation through microscope is based on viewing various cross sections and making a hand draft of what the observer has seen, analyzing and creating a record. Comparison of test sample image with reference not only requires an

experienced but is subjective and prone to human errors. By applying advanced technique of image processing and utilizing the capabilities of the recent advanced computing and data image storage facilities and the use of computer techniques for analyzing the shape and vein structure [4].

Image segmentation is the method of image processing for recognizing object. Generally, to recognize the object it is base requirement to eliminate the noise for doing this image preprocessing techniques plays important role. Image processing techniques allows extracting region of interest. These methods are important to enhance the image after actual segmentation of the image for object recognition for better effect. Image segmentation is the process of subdividing a digital image into various parts of the image in such a way that may form some of the meaningful regions for performing certain task for any application. The segmentation is nothing but extracting parts of image that collectively form the entire image. For the leaf image segmentation here it extracts the leaf components such as leaf shape, venation and/or texture etc. This process includes frequently used technique in image processing i.e. edge detection [3][4]. Plants are mostly identified by taxonomists and the process is usually lengthy. Leaves play an important role in plant identification as they are available nearly throughout the year. It is easy to carry them, and access and process through computerized methods. Plants are easily identified by visually recognizing the characteristics of leaves of the plants. Shape and venation pattern of the leaves are important characteristics in plant identification system [1]. Other leaf features like, complexity, blade shape, size, leaf apex, base, margin, texture and venation pattern play an important role in identification of plant.

Processing Steps

Plant classification system can perform following steps,

1. Plant leaf image acquisition
2. Leaf image preprocessing
3. Leaf image feature extraction
4. Classification

Plant Leaf Image Acquisition

Leaf image can be acquired by using camera by capturing the image or it can be acquired by using scanner. Figure 1 a) shows an original such sample image of the leaf which is taken from the ICL database.

Leaf Image Preprocessing

This step is processing the leaf image acquired this can be achieved by using the various morphological operations that remove the unwanted features of the leaf. By using morphological operation shape and vein features can be collected.

Figure 1 shows the leaf image in original, preprocessed to get in uniform formats for finding the leaf shape and venation features. The original green colored leaf image is converted into gray level format and then its leaf shape and venation pattern are extracted by using morphological operations.

(a) Original Image

(b) Gray level

(c)Shape and venation feature

Figure 1: Plant Leaf Image

Leaf Features Extraction

Leaf shape and vein features are extracted at this step. These features include geometrical features and morphological features.

a. Geometrical:

Diameter: The diameter is the distance between two points at margin of leaf image which is longest.

Length: The length of the major axis of the ellipse that is normalized to the leaf shape. (L)

Width: The length of the minor axis of the ellipse that is normalized to the leaf shape.(W)

Leaf Area: The total number of pixels in the region of leaf shape.

Leaf Perimeter: It is total number of pixels at leaf margin.[4]

b. Morphological features:

Smooth Factor: This the ratio of smoothing the image by using rectangular averaging filters with 5 × 5 rectangular averaging and 2 × 2 rectangular averaging.

Aspect Ratio: This is the ratio of length and width leaf image.

Form Factor: This indicates difference between leaf shape and a circle. This is obtained by the formula $4\pi A/P2$, where A is the leaf area and P is the perimeter of the leaf margin.[4]

Rectangularity: Rectangularity describes the similarity between a leaf and a rectangle. It is defined as L*W/A i.e. physiological length L into width W by area A.[4]

Narrow Factor: Narrow factor is defined as the ratio of the diameter D and length of leaf. [4]

Perimeter-Diameter Ratio: The ratio of leaf perimeter to leaf diameter.

Perimeter-Length and Width: This feature is defined as the ratio of leaf perimeter P and the sum of length L and width W, thus P/(L +W).[4]

Vein Features: By performing morphological opening with disk shaped structuring on gray scale leaf image vein features are extracted. For five results of vein extraction areas are obtained. For each of these vein features are obtained Av1/A, Av2/A, Av3/A, Av4/A, Av4/Av1 [4] where AV1 to AV5 are areas of vein results and A is the area of leaf.

Classification

This step performs classification. Classification step classify the plants based on leaf features acquired by following the morphological operation and image processing methods.

Artificial neural network: Artificial neural network is model for classification that classifies the plant species with their leaf features. This model takes leaf features as an input and generates the output that indicates the class to which plant species belongs. This model has input layer, hidden layer and output layer. Input layer take leaf features and hidden layer process that features and output layer generate the class specifications for the input features of the plant species as per their leaf features. Hidden layer perform many different operation with the activation functions.

Probabilistic Neural Network

Probabilistic neural network PNN is derived from Radial Basis Function Network (RBFNN) which is an artificial neural network using RBF. PNN has training speed many times faster than a back propagation BP network.[8] Probabilistic neural networks can be used for classification problems. When an input is presented, the Radial Basis Layer computes distances from the input vector to the training input vectors and produces a vector whose elements indicate how close the input is to a training input. The Competitive Layer sums these contributions for each class of inputs to produce as its net output a vector of probabilities. Finally, on the output of the second layer picks the maximum of these probabilities, and produces a 1 for that class and a 0 for the other classes. The architecture is shown, for this system in figure 2.

Working of the network:

The network classifies input vector Q into a specific class K because that class has the maximum probability to be correct. This PNN model has three layers: the Input layer, Radial

Basis Layer and the Competitive Layer. At Radial Basis layer Q input is transposed to P□ then evaluates vector distances between input vector and row weight vectors in weight matrix. These distances are scaled by Radial Basis Function nonlinearly. An input vector close to a training vector is represented by a number close to 1 in the output vector a1. If an input is close to several training vectors of a single class, it is represented by several elements of a1 that are close to 1. Then the Competitive Layer weights are set to the matrix T of target vectors. Each vector has a 1 only in the row associated with that particular class of input, and 0s elsewhere. The multiplication Ta1 sums the elements of a1 due to each of the K input classes. Finally, it produces a 1 corresponding to the largest element of d, and 0s elsewhere. Thus, the network classifies the input vector into a specific K class because that class has the maximum probability of being correct.

Figure2 shows the structure of the PNN. Table 1 shows the accuracy of various neural network models as reviewed for the plant classification.

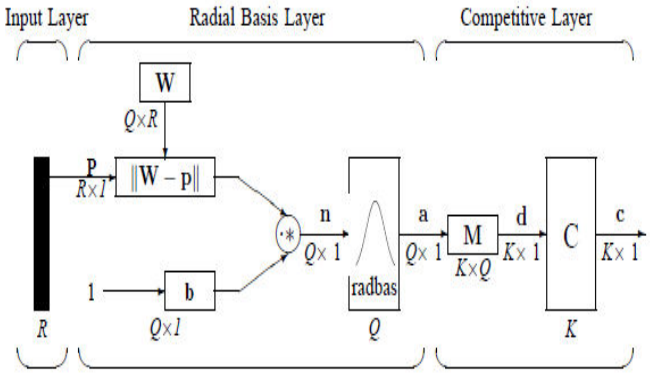

Figure. 2 :PNN Structure

Table 1: Accuracy of Various Methods

Scheme	Accuracy
MLNN	94%
1-NN	93%
BPNN	92%
k-NN (k=4)	92%
RBPNN	91%
k-NN (k=5)	86%

Source: A Leaf Recognition Algorithm for Plant Classification Using Probabilistic Neural Network(2007)

Conclusion

This review study shows the plants can be recognized and classified by using various image processing techniques and neural network approaches. With the help of leaf images of the plants they can be recognized and classified, which helps the rural people to easily recognize, identify and classify the plant species. This helps to avoid unnecessary death of many plant species.

References

[1] Wang Zhi, Zheru Chi and David Feng. 2003. Shape based leaf image retrieval. Vision, Image and Signal Processing, IEE Proceedings-. Vol. 150. No. 1. IET.

[2] Jens Langner. 2001. Leaves Recognition v.1.0, http://www.jens-langner.de/lrecog.

[3] Ji-Xiang Du, De-Shuang Huang, Xiao-Feng Wang and Xiao Gu. 2006. Computer-aided plant species identification (CAPSI) based on leaf shape matching technique, Transactions of the Institute of Measurement and Control.

[4] Stephen Gang Wu, Forrest Sheng Bao, Eric You Xu, Yu-Xuan Wang, Yi-Fan Chang and Qiao-Liang Xiang. 2007. A Leaf Recognition Algorithm for Plant Classification Using Probabilistic Neural Network", IEEE International Symposium on Signal Processing and Information Technology.

[5] Sumathi C.S. and A.V. Senthil Kumar, "Edge and Texture Fusion for Plant leaf classification", International journal of computer science and telecommunications (IJCST), vol3,Issue 6,June 2012.

[6] J.Reddy L. Manjunatha Rao, and M. N. Nachappa, "A study on multimodal biometric systems and Taxonomic identification of plant species", International Journal of Advanced Computer and Mathematical Sciences ISSN 2230-9624. Vol 3, Issue 2, 2012, pp 250-256 online available http://bipublication.com

[7] Abdul Kadir, Lukito Edi Nugroho, Adhi Susanto, Paulus Insap Santosa. 2011. Leaf Classification Using Shape, Color, and Texture Features, International Journal of Computer Trends and Technology.

[8] K.singh, I gupta , S Gupta, "SVM-BDT PNN and Fourier Moment Technique for Classification of Leaf Shape", International Journal of Signal Processing, Image Processing and Pattern Recognition Vol. 3, No. 4, December, 2010

[9] Gurpreet Kaur and Gurpinder Kaur. 2012. Classification of Biological Species Based on Leaf Architecture, International Journal of Engineering Research and Development.

[10] ManishaAmlekar, Dr. Ramesh R Manza,Dr. Pravin Yannawar, Dr. Ashok T Gaikwad, “Leaf Features Based Plant Classification Using Artificial Neural Network” , IBMRD's Journal of Management and Research, Print ISSN: 2277-7830, Online ISSN: 2348-5922