Classification of Leaf Using Geometric Features

Nidheesh P, Aishwarya Rajeev, Nikesh P

(Asst professor CIT ,Ponnampet[1] Asst professor CIT ,Ponnampet[2] Asst professor MIT ,Kannur[3])

Mail: nidheesh2005@gmail.com Contact:09916357357

Abstract— — Every leaf has its own identity and pocess some information that will help humans to identify and classify the plant by analyzing its leaves. The shape of leaf is a significant feature that most people use to recognise and classify a plant. It uses different parameters like diameter, physiological length, physiological width, leaf area and perimeter are basic geometry information can be extract from the leaf shape. The leaf identification is essential for scientists working in the agricultural and environmental fields. This work is a study of leaf identification and recognition system.The steps include the capturing the leaf image followed by a preprocessing. Later discussing about different classifiers and their accuracies

Keywords— General Regression Neural Network, Geometric, k-nearest neighbor algorithm, Morphological, Probabilistic Neural Network, Radial Basis Probabilistic Neural Network, recognition, Wavelet Transform.

1. INTRODUCTION

Plants have important functions in various areas, such as food, medical science, industry, and the environment. The classification of plants based on leaf identification become a latest research trend now. Each leaf carries substantial information that can be used to identify and classify the origin or the type of plant. In order for a computer to recognize the two objects as being similar, algorithms need to be developed that can successfully provide for object identification regardless of where they are located in the image or whether they are scaled or rotated. In a surgeon's view, images like x-ray and ultrasound scanning play an important rule to diagnose diseases and this method has been proven reliable for years. The researchers are also using the same method as doctors to simulate the same principle to recognize a plant using high quality leaf images and complex mathematical formula for a processor to decide the category and type of plants. In particular, it is understood that the proper way to extract plant features involves plant recognition based on leaf images. There are two methods widely used for plant recognition based on leaf image, are leaf color based recognition and leaf shape based recognition. In the color based conventional study, a simple color similarity between two images can be measured by comparing their color histogram. And in the shape based conventional study, use region and contour-based simple features and features could be considered time domain data. The main advantage of our proposed method is that only contour features are used in the matching process, the method is scale invariant and computationally efficient.

2. PROPOSED SYSTEM

The following section illustrates the fundamental of recognition and classification process by computer using a leaf image in order to recognize and classify a plant. Many of the proposed approaches are based on recognizing and classifying method. Recognition process normally occurs during preprocessing stage, followed by the extraction process as shown in the figure. Classification is the process which look up into a database for the comparison of the leaf features. The major steps are discussed in the consecutive sub-sections.

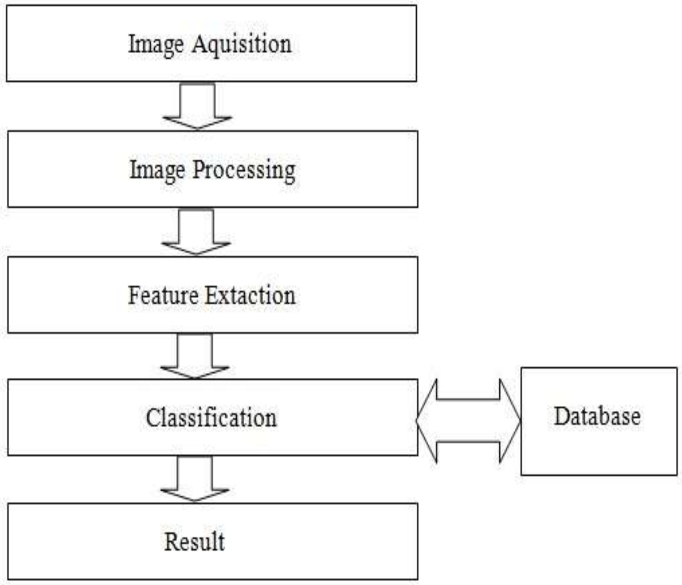

Figure 1. System description of plant recognition and classification

2.1 IMAGE AQUISITION

The images can be acquired using an image scanner or digital camera, even one captured using your mobile phone. There is no restriction on resolution and image format; The image can be an RGB image or a gray scale image. However, the image background needs be clean preferably white or any single colored with reasonable contrast with the leaf color and the leafstalk should be removed prior to image acquisition.

The preprocessing is done to identify the leaf in an image and discarding all other information other than the leaf shape. This step includs Leaf contour extraction. The leaf contour extraction can be expalin as follows. Most leaves have generally green color, while the color of leaves is changed by season or environmental factors. The color change of leaf image can cause decline of recognition performance or non-recognized problem. The color converting process on input image is the first step for leaf contour extraction, and it can set foundation to improve recognition performance irrelevant to the leaf color change. We can convert the input color leaf image to gray scale image by using the following expression

Gray=0.299*R+0.587G+0.114*B

The converted gray scale leaf image is converted to a binary image once again. The threshold conversion is performed as follows

$$\mathbf{B}(x,y) = \begin{cases} 0 & \text{if } f(x,y) \leq T \\ 255 & \text{if } f(x,y) > T \end{cases}$$

Where B(x,y) and f(x,y) are the intensity values of the gray scale image and the binary image, respectively, at position (x,y) , and T is the threshold value . Figure 2 shows an example of leaf contour extraction.

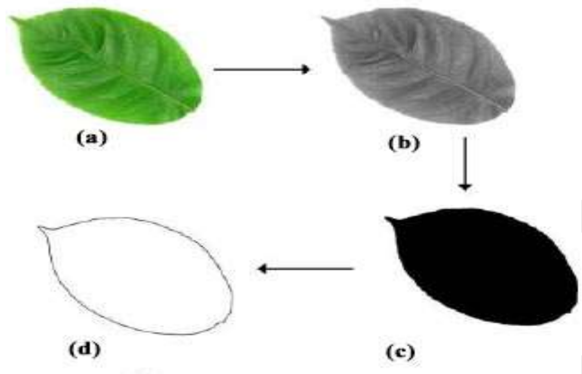

Figure 2. leaf contour extraction

2.2 IMAGE PREPROCESSING

The preprocessing is done to identify the leaf in an image and discarding all other information other than the leaf shape. This can be done with a little help from the user. The user can help identify the base-point and some reference points of the leaf. The leaf is extracted from the background and a binary image is produced where the background pixels are set to 0 or black and the pixels within the leaf is set to 1 or white. The remaining black pixels within the leaf blade are removed to produce an enhanced binary image. Then the tip of the leaf is located by finding out the furthest point (which is, in most cases, the tip of the leaf) from the base-point (selected by the user).

2.3 FEATURE EXTRACTION

After the pre-processing is done, feature extraction is easy. Our method takes into account only the shape of the leaf and the geometric features of the leaf.

Leaf feature using Geometric and digital morphological features

We describe geometric and digital morphological features in order to leaf feature extraction. We extract four basic geometric features as leaf length, leaf width, leaf area, leaf perimeter. The leaf length is defined as the longest distance between the centroid and the two ends on the margin of the leaf on opposite sides of the centroid. It is denoted by *LL* . The leaf width is defined as the distance between the intersection point with *LL* at the centroid and its opposite side on the margin of the leaf. It is denoted by *LW* . Figure 5 shows the procedure for obtaining *LL* and *LW* . The leaf area is the number of pixels in the leaf region. It is denoted by *LA*. The leaf perimeter is the number of pixels in the leaf contour. It is denoted by *LP*. We extract ten features based on digital morphological features using four basic geometric features and the study conducted as aspect ratio, form factor, rectangularity, perimeter ratio of the leaf length, perimeter ratio of the leaf length and leaf width, and five vein features. The aspect ratio is calculated using the leaf length *LL* and leaf width *LW* . It is defined as LL/ LW . The form factor is used to describe the difference between a leaf and a circle. It is defined as $4\pi LA/ LP^2$, where *LA* is the leaf area and *LP* is the perimeter of the leaf margin. The rectangularity describes the similarity between a leaf and a rectangle. It is defined as $LL*LW/ LA$, where *LL* is the leaf length, *LW* is the leaf width and *LA* is the leaf area. The ratio of perimeter to leaf length, representing the ratio of the leaf perimeter *LP* and leaf length *LL* , is calculated by LL/ LP . The perimeter ratio of the leaf length and leaf width is defined as the ratio of the leaf perimeter *LP* and the sum of the leaf length *LL* and leaf width *LW* , thus $LP/(LL+ LW)$. The Vein features have been extracted using the methods and the morphological openings on the gray scale images. The five features are as follows: Lv 1/A , Lv 2/A , Lv 3/A , Lv 4/ A, L*v4/Lv1* .

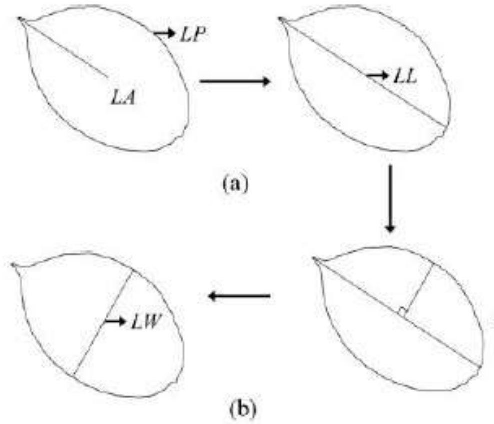

Figure 3. Procedure for Obtaining Leaf Length and Leaf Width: (a) leaf length,(b) leaf width

2.4 Classification

A classification based on different classifiers are implemented and compared. In the NN (Nearest Neighbour) classifier new sample is classified by calculating the distance to the nearest training case; the sign of that point then determines the classification of the sample. The k-nearest neighbor algorithm (K-NN) is a method for classifying objects based on closest training examples in the feature space. The k-NN classifier extends this idea by taking the k nearest points and assigning the sign of the majority. Larger k values help reduce the effects of noisy points within the training data set, and the choice of k is often performed through cross-validation .In the distance measurement based classification feature values of one class is added together then the average is found and thus form a core vector for one class. Similarly core vectors are finding for each class. Now this core vectors will represent the classes. For the testing purpose we will compare the features of each sample with these core vectors and the sample will go to the class which is having higher similarity with the values of testing data. Probabilistic Neural Network (PNN) to train the extracted values of 1800 leaves that was used and classified into 32 species of plant . The result is on average 90.312% accuracy. The testing to the proposed approach was conducted also with other general-purpose classification algorithms and it was found out that the algorithms only focused on leaf shape information. In other words, the proposed approach has an advantage because the approach is not only concentrating on leaf shape information in order to classify the plants. PNN method consists of several layers and the input layer will retrieve the vector values from the extraction process for training the method. However, colour and texture features become additional input to train the method. Consequently it is shown that there is an improvement of 3.44% in accuracy of the plant classification compared to 90.312% previously using algorithms. One of the neural network methods are known as feed-forward back-propagation neural network In this the number of nodes of input layer is the same as the number of extracted features and similarly with the output layer is same, as the number of plant categories, become the main reason why the method has been used for recognition purpose. Furthermore, the method that has been structured consists of three layers, which are 16 nodes of input layer, 32 nodes of hidden layer and 6 nodes of the output layer. The method has been trained based on 1200 samples which consist of 6 species of plant and 30 leaf images from each species. The result of the training was recorded based on the species. Still neural network algorithm has been chosen for recognition purpose. This time many were used Nearest Neighbour classifier (1-NN), k-Nearest Neighbour classifier (k-NN) and Radial Basis Probabilistic Neural Network (RBPNN) methods to train the samples. The vector values were retrieved from the previous segmentation process, where the algorithm was proposed to integrate with Wavelet Transform (WT) and Gaussian interpolation methods. As a result, it was 93.17% for (1-NN), 85.47% for (k-NN where k = 5) and 91.18% for (RBPNN). The finding is increasing the value of k will improve the stability of the proposed method in order to recognise the plant. General Regression Neural Network (GRNN) was also used for recognition purposes. Similarly with other methods, the vector values from the extraction process are input into the classifier to be trained. 10 species and 10 samples from each species have been used for training. The result from the testing is 100% accuracy rate of plant recognition and classification. Not only that any changes in the spread parameters of the GRNN will not affect the process of leaf recognition.

TABLE 1
LIST OF CLASSIFIER AND THE SUCCESSFULNESS RATE OF CLASSIFYING AND RECOGNIZING

Classifier	Accuracy
Probabilistic Neural Network (PNN)	90.312%
Nearest Neighbour (1-NN)	93.170 %
k -Nearest Neighbour (k-NN)	85.470 %
Radial Basis Probabilistic Neural Network (RBPNN)	91.180 %
General Regression Neural Network (GRNN)	100.000%

The extraction process and the extracted features also affect the classification rate. The reason is a different extraction methods will extract different features. Besides, more features are extracted and considered for recognition process, hence, more accurate the classification output. In fact, the same classifier has been used for recognition and classification, but because of different extraction methods were used earlier, therefore the accuracy rates are not the same as what has been produced.

3. CONCLUSION

The findings of our study are the types of leaf features that should be extracted, external factors that must be considered before the extraction process, types of extraction and classification methods that can be used for plant recognition and classification. In other words, the results of this study can be used as a specification of leaf features that must be considered for plant recognition and classification purposes as shown in **Table 1**. Finally, we can select some classifiers for testing and future development. The selection will be based on type of leaf features that can be extracted and recognised and ability of the pre-processing method to handle the noise or other external factors in the image.

In the future work, the selected classifiers will be tested based on the dataset and the results will be recorded. Only the better classifier will be used in future research work. However, we may have to consider images that contain many leaves in order to test the ability of the classifiers.

4. ACKNOWLEDGMENT

We wish to thank all our friends for supporting to complete our paper.

REFERENCES:

[1]. Mark S. Hammel, Przemyslaw Prusinkiewicz, and Brian Wyvill “Modelling Compound Leaves Using Implicit Contours” , Department of Computer Science University of Calgary, Alberta, Canada T2N 1N4June 1992. Proceedings of Computer Graphics International ’92 (Tokyo, Japan, 22–26 June 1992).

[2]. Z. Wang, Z. Chi and D. Feng “Shape based leaf image retrieval” @ IEE. 2003 Paper first

received 19th December **2001** and in revised form 16th September 2002

[3]. Yan Li, Zheru Chi, Member, IEEE, and David D. Feng, Fellow, IEEE” Leaf Vein Extraction

Using Independent Component Analysis” 2006 IEEE Conference on Systems, Man, and Cybernetics October 8-11, 2006, Taipei, Taiwan.

[4]. Stephen Gang Wu1, Forrest Sheng Bao2, Eric You Xu3, Yu-Xuan Wang4, Yi-Fan Chang5 and Qiao-Liang Xiang4 1 Institute of Applied Chemistry, Chinese Academy of Science, P. R. China 2 Dept. of Computer Science, Texas Tech University, USA “A Leaf Recognition Algorithm for Plant Classification Using Probabilistic Neural Network”, 29 Jul 2007.

[5] K. B. Lee and K. S. Hong, "Advanced Leaf Recognition based on Leaf Contour and Centroid for Plant Classification", The 2012 International Conference on Information Science and Technology, **(2012)** Shanghai, China.

[6] J. H. Kim, R. G. Huang, S. H. Jin and K. S. Hong, "Mobile-based flower recognition system", 2009 Third International Symposium on Intelligent Information Technology Application , **(2009)**, NanChang, China.

[4] T. H. Hsu, C. H. Lee and L. H. Chen, "An interactive flower image recognition system", Multimed. Tools Appl., vol. 53, no. 1, **(2011)**.

[7] S. J. Kim, B. W. Kim and D. P. Kim, "Tree Recognition for Landscape Using by Combination of Features of its Leaf, Flower and Bark", The SICE Annual Conference 2012, **(2011)** Tokyo, Japan.

[8] A. W. M. Smeulders, M. Worring, S. Santini, A. Gupta and R. Jain, "Content-based image retrieval at the end of the early years", IEEE Trans. Pattern Anal. Mach., vol. 22, **(2000)**.

[9] X. Wang, D. Huang, J. Dua, H. Xu and L. Heutte, "Classification of plant leaf images with complicated background", Appl. Math. Comput., vol. 205, **(2008)**.

[10] J. Du, D. Huang, X. Wang and X. Gu, "Computer-aided plant species identification (CAPSI) based on leaf shape matching technique", Trans. Inst. Meas. Control, vol. 23, **(2006)**.

[11] S. Zhang and Y. K. Lei, "Modified locally linear discriminant embedding for plant leaf recognition", Neurocomputing, vol. 74, (2011).