www.ioirp.com
International Journal of Innovative Research in Technology, Science & Engineering (IJIRTSE)
ISSN: 2395-5619, Volume – 1, Issue – 10. December 2015

DIGITAL IMAGE PROCESSING APPROACH FOR LEAF IDENTIFICATION

Suganthi P
PG Scholar, Department of EIE
Bannari Amman Institute of Technology
Sathyamangalm, Erode,India
suganthi.bit@gmail.com

Rajasekar L
Assistant Professor, Department of EIE
Bannari Amman Institute of Technology
Sathyamangalm, Erode,India
lrajsekar@gmail.com

ABSTRACT: ***Living plant identification based on images of leaf is a very challenging task in the field of Digital image processing. There has been little work reported on leaf image processing and identification. In recent years, several researchers have dedicated their work to leaf characterization. The disadvantages in Artificial Neural Network(ANN) is overcome by using a new approach that combines a Principal Component Analysis (PCA) method and Probabilistic Neural Network (PNN) classifier is proposed to extract leaf veins. The result is to improving the performance of leaf identification system using PCA. The system involved combination of features derived from shape, vein, and color of leaf. PCA was incorporated to identify the system then the results were inputted to the classifier that used for PNN. The result shows that PCA can increase the accuracy of the leaf identification system.***

Keywords

Leaf identification, PCA, PNN classifier, ANN, Feature extraction.

I. INTRODUCTION

Plants play an important role in our environment. Without plants there will be no existence of the earth's ecology. But in recent days, many types of plants are at the risk of extinction. To protect plants and to catalogue various types of flora diversities, a plant database is an important step towards conservation of earth's biosphere. There are a huge number of plant species worldwide. To handle such volumes of information, development of a quick and efficient classification method has become an area of active research. In addition to the conservation aspect, recognition of plants is also necessary to utilize their medicinal properties and using them as sources of alternative energy sources like bio-fuel. There are several ways to recognize a plant, like flower, root, leaf, fruit etc. The sample leaves are taken from various places, plants and shape. The image is captured and further work is carried out. 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, texture, color, aspect ratio, vein structure, entropy, compactness and so on. In recent times the Principal Component Analysis and Probabilistic Neural Network techniques have been applied towards automated procedures of plant recognition.

This paper implements a leaf recognition algorithm using easy-to-extract features and high efficient recognition algorithm. The main improvements of this paper are on feature extraction and the classifier. All features are extracted from digital leaf image. Except one feature, all features can be extracted automatically. 12 features are orthogonalized by Principal Components Analysis (PCA). As to the classifier, the PNN is use for its fast speed and simple structure. The whole algorithm is easy-to-implement, using common approaches. The paper proposes a scheme for automated recognition of types of plant species by analyzing shape features from digital images of their leaves. PCA is incorporated to the identification system to convert the features into Gabor features and then the results were inputted to the classifier that used Probabilistic Neural Network. The leaf identification system was developed by using Probabilistic Neural network. PNN is a kind of neural networks that can learn fast from training data and guarantees to converge to an optimal classifier as the size of the representative training set increases.

PNN is an implementation of sophisticated underlying mathematical principles to transforms a number of possibly correlated variables into a smaller number of variables called PCA algorithm, in which the operations are organized into a multilayered feed-forward network with four layers: 1) input layer, 2) pattern layer, 3) summation layer, and 4) output layer. When an input is presented, the first 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 training dataset of leaves. Leaf feature extraction represents the first part of the identification process. A filter-based leaf feature extraction is proposed to obtain some feature vectors which provide optimal characterizations of the visual content of images. For this reason we choose the two-dimensional Gabor filtering, a widely used image processing tool for feature extraction.

Image Processing Toolbox provides a comprehensive set of reference-standard algorithms and graphical tools for image processing, analysis, visualization, and algorithm development for restore noisy or degraded images, enhance images for improved intelligibility, extract features and analyze shapes. The identification of leaves from photographs implies several steps, starting with image preprocessing, feature extraction, plant identification, matching and testing and finally obtaining the results implemented in MATLAB.

II. RELATED WORK

In this paper, the leaf recognition algorithm using is easy-to-extract features and high efficient recognition algorithm. The main improvements of this paper are on feature extraction and the classifier. All features are extracted from digital leaf image and simulated using MATLAB.

A. INTRODUCTION TO IMAGE PROCESSING

In imaging science, image processing is processing of images using mathematical operations by using any form of signal processing for which the input is an image, such as a photograph or video frame; the output of image processing may be either an image or a set of characteristics or parameters related to the image. Most image-processing techniques involve treating the image as a two-dimensional signal and applying standard signal-processing techniques to it. This processing can be understood as applying standard one-dimensional signal processing techniques to two-dimensional signals. Image processing is a very important subject, and finds applications in such fields as photography, satellite imaging, medical imaging, and image compression, just to name a few.

In the past, image processing was largely done using analog devices. However, as computers have become more powerful, processing shifted toward the digital domain. Like one-dimensional digital signal processing, digital image processing overcomes traditional analog "problems" such as noise, distortion during processing, inflexibility of system to change, and difficulty of implementation. The image processing technique is to implementing for feature extraction. As the board we have does not support a direct connection for the input image, we will use MATLAB to output the image as a matrix and store it in the data memory of the digital image processing. The parallel port connection is to get our input data into the board. The digital image processing will then do the processing and write the output data in the program memory. Then the extracted output data and go back to MATLAB to analyze the results.

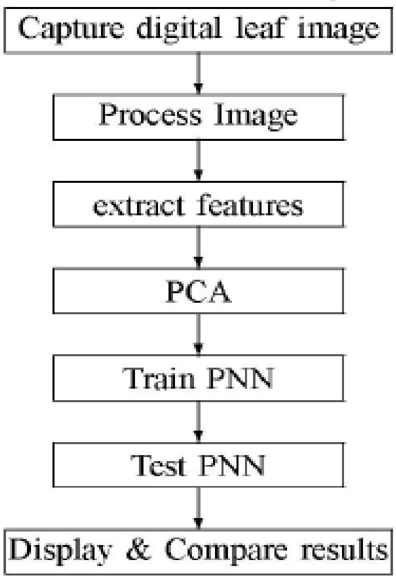

Fig 1 Flow Diagram of Proposed Scheme

The cameras and recording media available for modern digital image processing applications are changing at a significant pace. To dwell too long in this section on one major type of camera, such as the CCD camera, and to ignore developments in areas such as charge injection device (CID) cameras and CMOS cameras is to run the risk of obsolescence. Nevertheless, the techniques that are used to characterize the CCD camera remain "universal" and the presentation that follows is given in the context of modern CCD technology for purposes of illustration

B. IMAGE PRE-PROCESSING

1.CONVERTING RGB IMAGE TO BINARY IMAGE

The leaf image is acquired by scanners or digital cameras. Since they are not found any digitizing device to save the image in a lossless compression format, the image format here is JPEG. All leaf images are 800 x 600 resolutions. There is no restriction on the direction of leaves when photoing. An RGB image is firstly converted into a grayscale image.Eq.1 is the formula used to convert RGB value of a pixel into its grayscale value.

$$Gray=0.2989*R+0.5870*G+0.1140*B \qquad (3.1)$$

Where R,G,B correspond to the color of the pixel, respectively. The level to convert grayscale into binary image is determined according to the RGB histogram. The accumulate pixel values to color R,G,B respectively for 3000 leaves and divide them by 3000,the number of leaves. The average histogram to RGB of 3000 leaf images is shown as figure 2

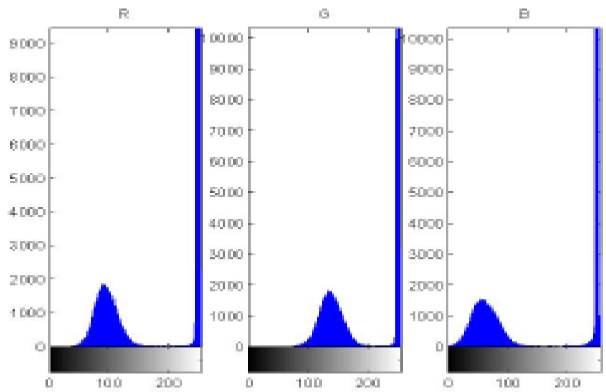

Fig 2 RGB histogram

There are two peaks in every color's histogram. The left peak refers to pixels consisting of the leaf while the right peak refers to pixels consisting of the white background. The lowest point between two peaks is around the value 242 on the average. So, choose the level as 0.95(242/255=0.949).The output image replaces all pixels in the input image with luminance greater than the level by the value 1 and replaces all other pixels by the value 0. A rectangular averaging filter of size 3 x 3 is applied to filter noises. Then pixel values are rounded to 0 or 1.

a) BOUNDARY ENHANCEMENT

When mentioning the leaf shape, the first thing appears in your mind might be the margin of a leaf. Convolving the image with a laplacian filter of following 3 x 3 spatial mask

International Journal of Innovative Research in Technology, Science & Engineering (IJIRTSE)
ISSN: 2395-5619, Volume – 1, Issue – 10. December 2015

$$\begin{matrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{matrix}$$

The margin of leaf image is existed. An example of image pre-processing is illustrated in Figure 3.To make boundary as a black curve on white background, the “0” “1” value of pixels is swapped.

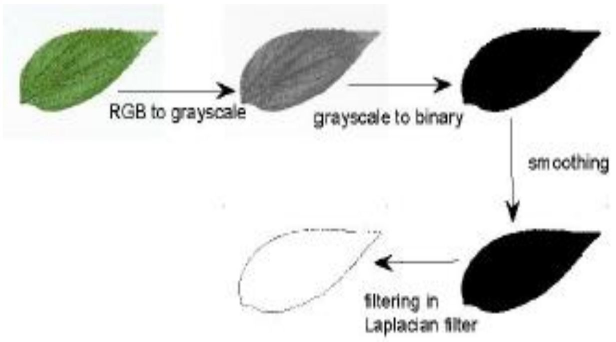

Fig 3 A pre-processing example

b) FEATURE EXTRACTION

Feature extraction starts from an initial set of measured data and builds derived values (features) intended to be informative, non redundant, facilitating the subsequent learning and generalization steps, in some cases leading to better human interpretations. Feature extraction is related to dimensionality reduction.

When the input data to an algorithm is too large to be processed and it is suspected to be redundant (e.g. the same measurement in both feet and meters, or the repetitiveness of images presented as pixels), then it can be transformed into a reduced set of features (also named a "features vector"). This process is called feature extraction. The extracted features are expected to contain the relevant information from the input data, so that the desired task can be performed by using this reduced representation instead of the complete initial data. In general it involves reducing the amount of resources required to describe a large set of data. When performing analysis of complex data one of the major problems stems from the number of variables involved. Analysis with a large number of variables generally requires a large amount of memory and computation power or a classification algorithm which over fits the training sample and generalizes poorly to new samples. Feature extraction is a general term for methods of constructing combinations of the variables to get around these problems while still describing the data with sufficient accuracy.
In this paper, 12 commonly used digital morphological features (DMFs), derived from 5 basic features, are extracted so that a computer can obtain feature values quickly and automatically (only one exception).

2. BASIC GEOMETRIC FEATURES

a) DIAMETER

The diameter is defined as the longest distance between any two points on the margin of leaf. It is denoted as D.

b) PHYSIOLOGICAL LENGTH

The only human interfered part of algorithm is needed to mark the two terminals of the main vein of the leaf via mouse click. The distance between the two terminals is defined as the physiological length. It is denote as Lp.

c) PHYSIOLOGICAL WIDTH

Drawing a line passing through the two terminals of the main vein, one can plot infinite lines orthogonal to that line. The number of intersection pairs between those lines and the leaf margin is also infinite .The longest distance between points of those intersection pairs is defined at the physiological width. It is denoted by Wp. Since the coordinates of pixels are discrete, we consider two lines are orthogonal if their degree is .The relationship between physiological length and physiological width is illustrated in Figure 3.

d) LEAF AREA

The values of leaf area are easy to evaluate, just counting the number of pixels of binary value 1 on smoothed leaf image. It is denoted by A.

e) LEAF PERIMETER

Denoted as P, leaf perimeter is calculated by counting the number of pixels consisting leaf margin.

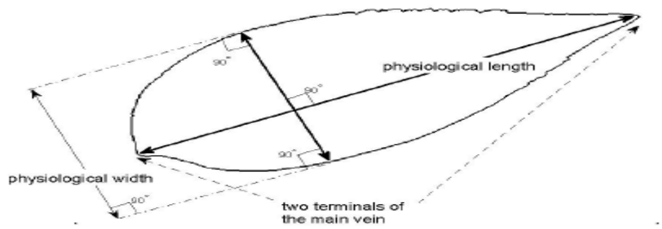

Fig 4 Relationship between physiological length and physiological width

f) VEIN FEATURES

Based on pixels on the vein, 2 vein features can be derived. In this research, vein was generated by using a morphological operation called opening [3]. That operation was performed on the gray scale image using disk-shaped structuring element of radius 1 and 2 and subtracted remained image by the margin. By using that operation, a structure like vein was obtained. Then, 2 features were calculated by using following formulas:

V_1

V1 and V2 represent features of the vein, A1 and A2 are total pixels of the vein, and A denotes total pixels on the part of the leaf.

C. DIGITAL MORPHOLOGICAL FEATURES

Based on 5 basic features introduced the previously. The 12 digital morphological features used for leaf recognition. Smooth Factor is the effect of noises to image area is use to describe the smoothness of leaf image. In this paper, smooth factor is defined as the ratio between areas of leaf image smoothed by 5 x 5 rectangular averaging filters. The aspect ratio is defined as the ratio of physiological length Lp to physiological width Wp, thus Lp/Wp. Form factor is used to describe the difference between a leaf and a circle. It is defined as $4\pi A/P^2$, where A is the leaf area and P is the perimeter of the leaf margin. Rectangularity describes the similarity between a leaf and a rectangle. It is defined as LpWp/A, where Lp is the physiological length, Wp is the physiological width and A is the leaf area. Narrow factor is defined as the ratio of the diameter D and physiological length, Lp, thus D/Lp. Ratio of perimeter to diameter, representing the ratio of leaf perimeter P and leaf diameter D ,is calculated by P/D. Perimeter ratio of physiological length and physiological width is the feature is defined as the ratio of leaf perimeter P and the sum of physiological length Lp and physiological width Wp, thus P/(Lp+Wp).Vein feature is to perform morphological opening on grayscale image with falt, disk-shaped structuring element of radius 1,2,3,4 and subtract remaining image by the margin. The result looks like the vein. That is why following 5 features are called vein features. Areas of left pixels are denoted as Av1, Av2, Av3 and Av4 respectively. Then the last 5 features are obtain they are Av1/A, Av2/A,Av3/A,Av4/A,Av4/Av1.Now the step of feature acquisition is finished go on to the data analysis section.

1) PCA USE FOR IMAGE COMPRESSION

Data volume reduction is a common task in image processing. There is a huge amount of algorithms [1, 2, 4] based on various principles leading to the image compression. Algorithms based on the image color reduction are mostly loss but their results are still acceptable for some applications. The image transformation from color to the gray-level (intensity) image I belongs to the most common algorithms. Its implementation is usually based on the weighted sum of three color components R, G, B according to relation

$$I = w1R + w2$$

The R, G and B matrices contain image color components; the weights were deter-mined with regards to the possibilities of human perception. The PCA method provides an alternative way to this method. The idea is based on Equation (4.17) where the matrix A is replaced by matrix Al in which only l largest (instead of n) eigen values are used for its forming. The vector xˆ of reconstructed variables is then given by relation.

$$\hat{X} = AkTy + m_x$$

True-color images of size M x N are usually saved in the three-dimensional matrix P with size M x N x 3 which means that the information about intensity of color components is stored in the 3 given planes. The vector of input variables x in Equation (4.17) can be formed as the n=3-dimensional vector of each color. Forming three 1-dimensional vectors x1,2,3 from each plane P(M, N, i) with the length of M.N can be advantageous for better understanding and programming. The covariance matrix Cx and corresponding matrix A are then evaluated and the 3-dimensional reconstructed vector xˆ according to Equation (4.18) can be called as the first, the second and the third component of the given image. The matrix theory implies that the image obtained by reconstruction with the matrix A1 (only the first - largest eigen value was used for its definition) contains the majority of information so this image should have the maximum contrast. These properties could be significant in the following image processing. There is a selected real picture P and its R, G, B components in the Fig 1. Its three reconstructed components obtained according to Equation (4.18) for each Eigen values are presented in Figure 4.2. The comparison of intensity images obtained from the original image as weighted color sum evaluated by Equation (4.17) and as the first principal component is presented in Figure 4.3. The Eigen values sorted in descending order belonging to the selected image.

To reduce the dimension of input vector of neural network, PCA is used to orthogonalize 12 features. The purpose of PCA is to present the information of original data as the linear combination of certain liner irrelevant variables. Mathematically, PCA transforms the data to a new coordinate system such that the greatest variance by any paperion of the data comes to lie on the first coordinate, the second greatest variance on the second coordinate, and so on. Each coordinate is called principal component. In this paper, the contribution of first 5 principal components is 93.6%. To balance the computational complexity and accuracy, adopt for 5 principal components. When using this algorithm, one can use the mapping f: to obtain the values of components in the new coordinate system and to implement.

2) INTRODUCTIONS TO PNN

The PNN is based on the theory of Bayesian classification and the estimation of probability density function (PDF). The idea of PNN was first introduced by Donald F. Specht in 1990. Because of ease of training and a sound statistical foundation in Bayesian estimation theory, PNN has become an effective tool for solving many classification problems. In fact, By replacing the sigmoid activation function often used in neural networks with an exponential function, a PNN that can compute nonlinear decision boundaries which approach the Bayes optimal is formed.PNN is closely related to Parzen window pdf estimator. A PNN consists of several sub-networks, each of which is a Parzen window pdf estimator for each of the classes. The development of the probabilistic neural network relies on Parzen windows classifiers. The Parzen windows method is a non-parametric procedure that synthesizes an estimate of a PDF by superposition of a number of windows, replicas of a function (often the Gaussian). The Parzen windows classifier takes a classification decision after calculating the probability density function of each class using the given training examples. The multicategory Classifier decision is expressed as follows:

$$p_k f_k > p_j f_j$$

Where p_k is the prior probability of occurrence of examples from class k and f_k is the estimated PDF of class k. It is called a "neural network" because of its natural mapping onto a two-layer feed forward network.

3) ARCHITECTURE OF PNN

The PNN architecture consists of four layers: input layer, pattern layer, summation layer, and decision layer.

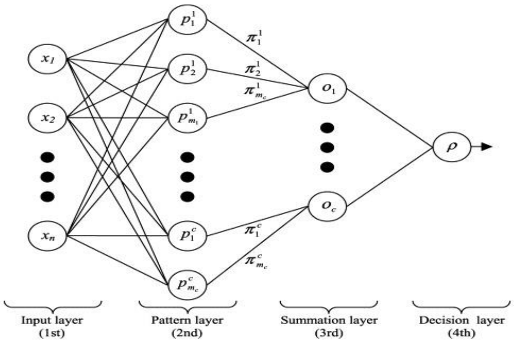

Fig 5 Architecture of PNN

The Figure 5.1 shows a PNN structure that recognizes c classes. The first layer shows the input pattern with n features. The number of nodes in the pattern layer is equal to the number of training instances. The number of nodes in the summation layer is equal to the number of classes in the training instances. The input layer is fully connected to the pattern layer. The input layer does not perform any computation and simply distributes the input to the neurons in the pattern layer. The pattern layer is semi-connected to the summation layer. Each group of training instances corresponding to each class is just connected to one node in the summation class. In other words, the summation units simply sum the inputs from the pattern units that correspond to the category from which the training pattern was selected.

4) PNN FOR TRAINED DATA

The PNN works by creating a set of multivariate probability densities that are derived from the training vectors presented to the network. The input instance with unknown category is propagated to the pattern layer. Once each node in the pattern layer receives the input, the output of the node will be computed

$$\pi_i{}^c = \frac{1}{(2\pi)^{n/2}\sigma^n} exp\left[-\frac{(x-x_{ij})^T(x-x_{ij})}{2\sigma^2}\right]$$

Where d is the number of features of the input instance x, σ is the smoothing parameter, and x_{ij} is a training instance corresponding to category c. The summation layer neurons compute the maximum likelihood of pattern being classified into c by summarizing and averaging the output of all neurons that belong to the same class

$$p_i(x) = \frac{1}{(2\pi)^{n/2}\sigma^n}\frac{1}{N_i}\sum_{t-1}^{N_i} exp\left[-\frac{(x-x_{ij})^T(x-x_{ij})}{2\sigma^2}\right]$$

Where N_i denotes the total number of samples in class c. If the a priori probabilities for each class are the same, and the losses associated with making an incorrect decision for each class are the same, the decision layer unit classifies the pattern x in accordance with the Bayes's decision rule based on the output of all the summation layer neurons

$$C(x) = \text{argmax}\ \{p_i(x)\},\ i = 1,2,\ldots,c$$

Where C(x) denotes the estimated class of the pattern x and m is the total number of classes in the training samples. If the a priori probabilities for each class are not the same, and the losses associated with making an incorrect decision for each class are different, the output of all the summation layer neurons will be

$$C(x) = \text{argmax}\ \{p_i(x)\ cost_i(x)\ apro_i(x)\},\ i = 1,2,\ldots,c$$

Where $cost_i(x)$ the cost is associated with misclassifying the input vector and $apro_i(x)$ is the prior probability of occurrence of patterns in class c. There is no iteration or computation of weights. For a large number of Gaussians in a sum, the error buildup can be significant. Thus the feature vectors in each class may be reduced by thinning those that are too close to another one and making σ larger. However, due to each pattern layer Gaussian component density p(x) being derived from one training vector, the PNN is limited to applications involving relatively small datasets. Large datasets would lead to large network architectures, which would have an adverse impact on computational complexity. In addition, this could saturate the feature space with overlapping Gaussian functions that would increase the rate of misclassification.

5) PNN CLASSIFIER

An artificial neural network (ANN) is an interconnected group of artificial neurons simulating the thinking process of human brain. One can consider an ANN as a "magical" black box trained to achieve expected intelligent process, against the input and output information stream. Thus, there is no need for a specified algorithm on how to identify different plants.PNN is derived from Radial Basis Function (RBF) Network which is an ANN using RBF, RBF is a bell shaped function that scales the variable nonlinearly. PNN is adopted for it has many advantages. Its training speed is many times faster than a BP network.PNN can approach a Bayes optimal result under certain easily met conditions. Additionally, it is robust to noise examples. This chooses it for its simple structure and training manner. The most important advantage of PNN is that training is easy and instantaneous. Weights are not "trained" but assigned. Existing weights will never be alternated but only new vectors are inserted into weight matrices when training. So it can be used in real-time. Since the training and running procedure can be implemented by matrix manipulation, the speed of PNN is very fast.

a) PNN LAYERS

The network classifies input vector into a specific class because that class has the maximum probability to be correct. In this paper, the PNN has three layers they are the input layer, Radial Basis layer and the Competitive layer. Radial Basis Layer evaluates vector distance between input vector and row weight vector in weight matrix. Then distance is scaled by Radial Basis Function nonlinearly. Then the Competitive Layer finds the shortest distance among them, and thus finds the training pattern closest to the input pattern based on their distance

b) NETWORK STRUCTURE

The network structure is illustrated in figure. The symbols and notations used in the book Neural Network Design .These symbols and notations are also used by MATLAB Neural Network Toolbox. Dimensions of arrays are marked under their names.

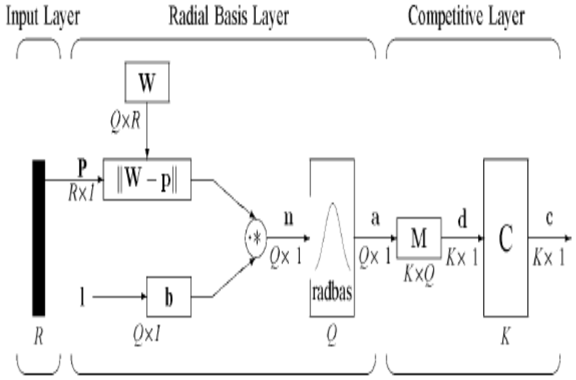

Fig 6 Network Structure

C) CHARACTERISTICS OF RADIAL BASIS LAYER

The *i*-th element of a equals to 1 if the input p is identical to the *i*-th row of input weight matrix W.A radial basis neuron with a weight vector close to the input vector p produces a value near 1 and then its output weights in the competitive layer will pass their values to the competitive layer will pass their values to the competitive function which will be discussed later. It is also possible that several training patterns.

d) COMPETITIVE LAYER

There is no bias in competitive Layer. In competitive Layer, the vector firstly multiplied with layer weight matrix M, producing an output vector d. The competitive function, denoted as c in figure, produces a 1 corresponding to the largest element of d, and 0's elsewhere. The output vector of competitive function is denoted as c. The index of 1 in c is the number of plants that the system can classify. It can be used as the index to look for the scientific name of this plant. The dimension of output vector, K, is 32.

e) NETWORK TRAINING

Totally 1800 pure leaves are sampled to train this network. Those leaves are common plants in china. Details about the leaf numbers of different kinds of plants are given in table. The reason why the samples are different pieces of leaves to different plants is the difficulty to sample leaves varies on plants.

f) RADIAL BASIS LAYER WEIGHTS

W is set to the transpose of R x Q matrix of Q training vectors. Each row of W consists of 5 principal variables of one training samples. Since 1800 samples are used for training, Q=1800.

g) RADIAL BASIS LAYER BIASES

All biases in radial basis layer are all set to resulting in radial basis functions that cross 0.5 at weighted inputs of is called the spread constant of PNN. The values of s cannot be selected arbitrarily. Each neuron in radial basis layer will respond with 0.5 or more to any input vectors within a vector distance of s from their weight vector. A too small s value can result in a solution that does not generalize from the input/target vectors used in the design. In contrast, if the spread constant is large values (near 1.0) for all the inputs used to design the network. In this paper, the s is set to 0.03 according to the experience.

h) COMPETITIVE LAYER WEIGHTS

M is set to K x Q matrix of Q target class vectors. The target class vectors are converted from class indices corresponding to input vectors. This process generates a sparse matrix of vectors, with one 1 in each column, as indicated by indices. For example, if the i-th sample in training set is the *j*-th kind of plant, then one 1 on the *j*-th row of *i*-th column of M.

D. EXPEREMENTAL RESULTS

To each kind of plant, 10 pieces of leaves from testing sets are used to test the accuracy of algorithm. Numbers incorrect recognition is listed in the last column of Table5.1.The average accuracy is 94.12%.Some species get a low accuracy in Table5.1.Due to the simplicity of algorithm framework, and add the more number of features to boost the accuracy.

Compared the accuracy of the algorithm with other general purpose (not only applicable to certain species) the accuracy of the algorithm is very similar to other schemes. Considering the advantage is respect to other automated/semi-automated general purpose schemes, easy-to-implement framework and fast speed of PNN, the performance is very good.

1) SIMULINK MODEL FOR PCA ALGORTHIM AND PNN CLASSIFIER

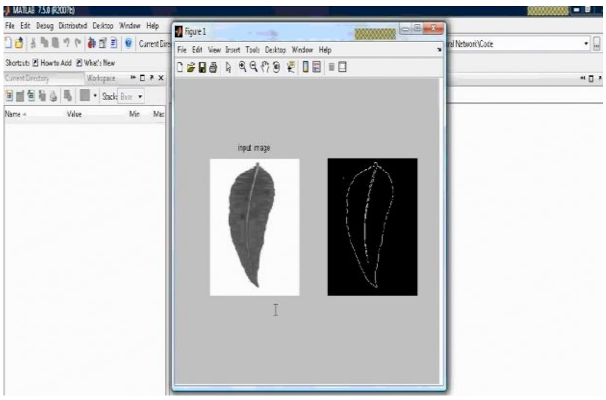
Fig 5 Simulations Window for PNN Classifier

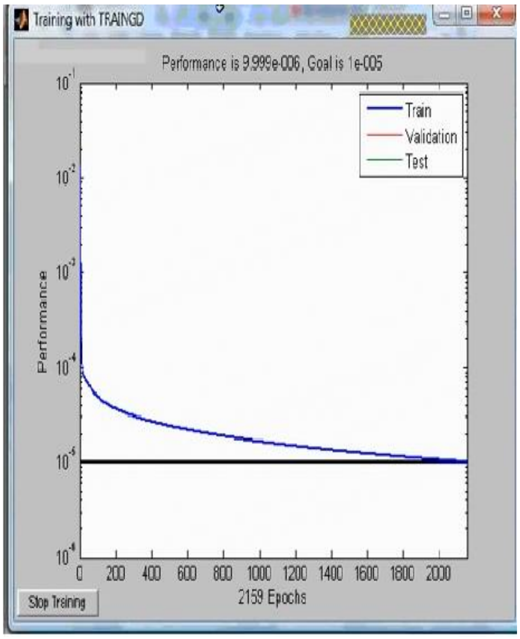

Fig 8 Test Result For PNN Trained Dataset

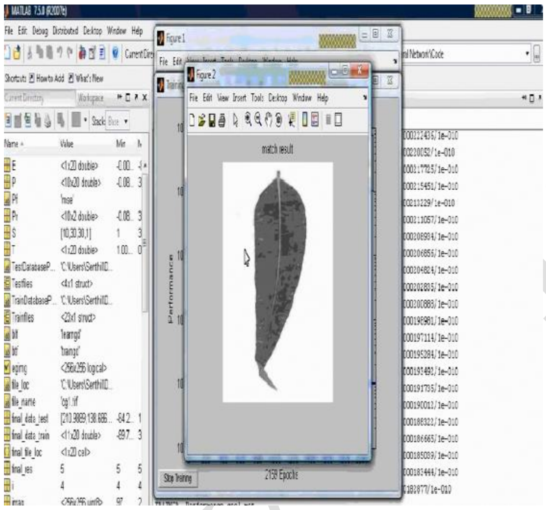

Fig 9 Simulation Window For Algorthim

E. CONCLUSION

This paper introduces a neural network approach for plant leaf recognition. The computer can automatically classify 32 kinds of plants via the leaf images loaded from digital cameras or scanners. PNN is adopted for it has fast speed on training and simple structure. 12 features are extracted and processed by PCA to form the input vector of PNN. Experimental result indicates that our algorithm is workable with accuracy greater than 94.12% on 32 kinds of plants. Compared with other methods, this algorithm is fast in execution, efficient in recognition and easy in implementation.

The future scope is to implement the PCA and PNN classifier in the real time process. In order to achieve high accuracy the more number of features are extracted and trained using PNN classifier in real time process.

REFERENCES

1. H. Fu and Z. Chi..,(2006) ‘Combined thresholding and neural network approach for vein pattern extraction from leaf images’, IEE Proc.-Vis. Image Signal Process., Vol. 153, No. 6.
2. Ji-Xiang Du, Xiao-Feng Wang and Guo-Jun Zhang..,(2007) ‘Leaf shape based plant species recognition’, Science Direct, Journal of Applied Mathematics and Computation 185.
3. Stephen Gang Wu1, 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 978-1-4244-1835.
4. PV Nageswara Rao, T Uma Devi, GR Sridhar and Allam Appa Rao..,(2005) ‘A Probabilistic Neural Network Approach For Protein Superfamily Classification’, Journal of Theoretical and Applied Information Technology.
5. Abdul Kadir, Lukito Edi Nugroho, Adhi Susanto and Paulus Insap Santosa..,(2010) ‘Performance Improvement of Leaf Identification System Using Principal Component Analysis’, International Journal of Advanced Science and Technology Vol. 44.
6. Abdolvahab Ehsanirad and Sharath Kumar Y. H..,(2010) ‘Leaf recognition for plant classification using GLCM and PCA methods’, Journal of Computer Science & Technology vol 33-34.
7. Jayamala K. Patil1 and Raj Kumar..,(2011) ‘Advances In Image Processing For Detection Of Plant Diseases’ Journal of Advanced Bioinformatics Applications and Research pp 135-14.
8. Pushpa Rani M.K.,(2015) ‘Detection and Analysis of Plant Leaf Diseases using Image Processing’, International Journal of Computer Applications (0975 – 8887).
9. Z. Wang, Z. Chi, D. Feng,,(2002) ‘ Fuzzy integral for leaf image retrieval’, Proc. Fuzzy Syst’, Journal of Computer Science & Technology 372–377.
10. Cholhong Im and H. Nishida.,(1999) ‘Recognizing plant species by normalized leaf shapes’, Journal of Advanced Bioinformatics Applications and Research pp397–404.