International Journal of Innovative Research in Advanced Engineering (IJIRAE) ISSN: 2349-2163
Issue 6, Volume 2 (June 2015) www.ijirae.com

Recognition of Plant Species based on leaf images using Multilayer Feed Forward Neural Network

Sapna Sharma[#1], Dr. Chitvan Gupta[#2]
CSE Department, Noida Institute of Engineering and Technology, Gr.Noida, Uttar Pradesh, INDIA-201306

Abstract- It is a challenging task to analyze plant leaf images by a layman because there are very minute variations in some plant leaf images & larger data set for analysis. It is a quite difficult to develop an automated recognition system which could process on a large information and provide a correct estimation. Artificial neural networks have been successfully applied to problems in pattern recognition, classification and image analysis. In this paper, Multilayer feed- forward networks are trained using back propagation learning algorithm. The main objective of this paper is to develop a classification system for agriculture and ayurvedic plants by image pre-processing, leaf contour, feature extraction, network training and classification. In our proposed algorithm we tested 440 leaves of different 16 classes and the method yields accuracy greater than 90%.

Keywords- Plant leaf classification and recognition, Image pre-processing, Artificial neural network, Back-propagation algorithm, Training, Testing and PCA

I. INTRODUCTION

Plants are the most important forms of life on our planet. Plants maintain the balance of O_2 and CO_2 of earth's atmosphere. In addition, plants are important means of livelihood and production of human beings. Computerized plant identification system can be very helpful in botanical garden or natural reserve park management, new plant species discovery, plant taxonomy, exotic plant detection and so on. A computer based plant identification or classification system can use different characteristics of the flora, starting at very simple level such as: shape and color of the leaf, flower and fruit type, branching style, root type, seasonality, outlook to very complex such as cell and tissue structure, genetic structure.

A substantial amount of work has been done on leaf shape based plant classification and recognition. Wu et al. [1] extracted 12 commonly used digital morphological features which were orthogonalized into 5 principal variables using PCA [24] They used 1800 leaves to classify 32 kinds of plants using probabilistic neural network system [23][15]. Wang et al. [2] employed centroid contour distance (CCD) curve, eccentricity and angle code histogram (ACH). He show that the proposed approach can achieve a better retrieval performance than both the curvature scale space (CSS) method and the modified Fourier descriptor (MFD) method. Fu et al. [3] also used centroid-contour distance (CCD) curve to represent leaf shapes in which an ontology-based leaf classification system is proposed, where in machine learning techniques play a crucial role.

Du et al. [4] an efficient computer-aided plant species identification (CAPSI) approach is proposed, which is based on plant leaf images using a shape matching technique. Gu et al. [5] used the result of segmentation of leaf's skeleton based on the combination of wavelet transform (WT) and Gaussian interpolation. Wang et al. [6] extracted several geometric features like rectangularity, circularity, eccentricity and seven moment invariants for classification. He introduces a method of recognizing leaf images based on shape features using a hyper sphere classifier. Some [1][3][7] approaches employed artificial neural network for its fast performance. Others [5][6] employed k-nearest neighbour (k-NN) classifier to classify plants. Du et al. [7] introduced shape recognition based on radial basis probabilistic neural network which is trained by orthogonal least square algorithm (OLSA) and optimized by recursive OLSA.

Jan Fusser [19] proposed a new set of moment invariants with respect to rotation, translation, and scaling suitable for recognition of objects having -fold rotation symmetry. Moment invariants described earlier cannot be used for this purpose because most moments of symmetric objects vanish. Du and Zhang et al. [13] approach to a new classification method, named as move median centers (MMC) hypersphere classifier based on digital morphological feature is proposed. An effective technique for plant species recognition by means of leaf image is recommended by Hossain and Amin et al. [20]. This technique executes only on the plants with wide flat leaves which are more or less two dimensional in general. A new technique for feature extraction from natural image like plant leaf is developed by Prasad et al. [21] for automated living plant species identification which would be helpful for botanical students to carry out their research for plant species identification. Abdul kadir et al. [22] builds a foliage plant identification system for 60 kinds of leaves. In this case, Zernike moments were combined with other features: geometric features, color moments and gray-level co-occurrence matrix (GLCM). A. Kadir et al. [25] reported the results of experiments in improving performance of leaf identification system using Principal Component Analysis (PCA) to convert the features into orthogonal features and then the results were inputted to the classifier that used PNN [23]. This approach has been tested on 2 datasets, Foliage and Flavia which contain various color leaves (foliage plants) and green leaves respectively. The results showed that PCA [24] can increase the accuracy of the leaf identification system on both datasets. Rashad et al. [26] introduced an approach to plant classification based on characterization of texture properties.

Anant [16] gives the brief description of various plant recognition techniques and Anant et al. [17] gives an approach to classify plant leaf images on the basis of higher order moment invariants and texture analysis using nearest neighbour classifier.This paper presents a plant leaf classification technique using neural network approaches with its back propagation algorithm. The whole algorithm is explained here section wise as shown in Figure 1. Section II discusses image pre-processing. Section III introduces the features which are essential for classification and decides the parameters of different classes of plants. Section IV explains the artificial neural network structure and Section V discusses the result of our work. Last Section VI concludes this paper.

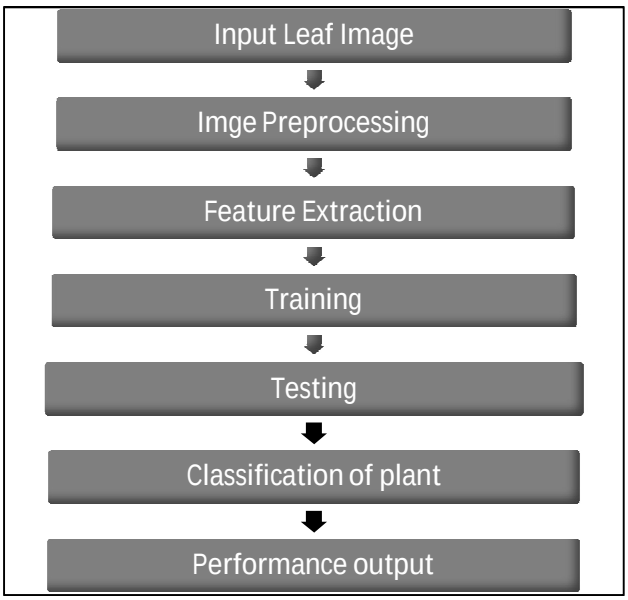


Figure 1.Stages in the Proposed Model

II. IMAGE PREPROCESSING

The leaf images can be acquired using a digital camera which is also embedded in our cell-phone. There is no restriction on resolution and image format. Generally the scanned image or digital image in two dimensional in nature and RGB image. In some cases we can use gray images also. However, the image background [12] needs be clean preferably white or any single colored with reasonable contrast with the leaf color and there is no leafstalk. This can be done by segmentation so that we process only the leaf shape.

The main goal of preprocessing is to identify the leaf in an image and discarding all other information other than the leaf shape. The first step is resizing all the database leaves in 256×256 and then converting RGB image to grayscale image. One of the main tasks of this application is the detection of specific tokens in a leaf image. These tokens will then be the basis of the neuronal network calculations.

Convolving the image with a Laplacian filter of 3 × 3 spatial mask or Prewitt edge detection produces an image where higher greylevel values indicate the presence of an edge between two objects. The Prewitt Edge Detection filter computes the root mean square of two 3x3 templates. It is one of the most popular 3x3 edge detection filters. To make boundary as a black curve on white background, the “0” “1” value of pixels is swapped. Thus we can find the edges of leaf i.e. respective contour of the leaf images as shown in Figure 2.

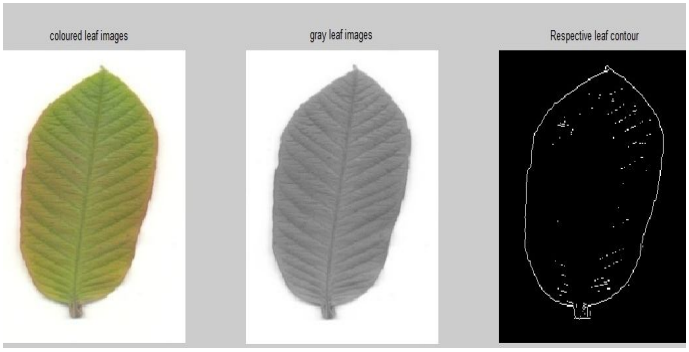


Figure 2. An example of leaf image pre-processing

III. FEATURE EXTRACTION

This approach uses 12 digital and morphological features which are briefly explained as:

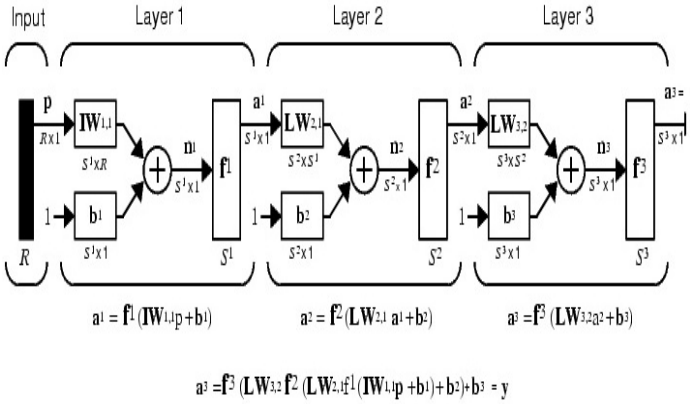
Major axis length: The line segment connecting the base and the tip of the leaf is the major axis.
Minor axis length: The maximum width, which is perpendicular to the major axis, is the minor axis of a leaf.
Area: Area is the actual number of pixels in the region. The area of leaf in a preprocessed image is the number of white or '1' pixels.
Convex area: It specifies the number of pixels in 'Convex Image'.
Filled area: The total number of on pixels in Filled Image is known as Filled Area.
Eccentricity: The eccentricity is the ratio of the distance between the foci of the ellipse and its major axis length. The value is between 0 and 1. The eccentricity is defined as Eccentricity = w/l where w is the length of the minor axis and l is the length of the leaf' major axis. This feature can be used to differentiate the rounded leaf and the long one.
Perimeter: The distance around the boundary of the region is called perimeter. The command 'regionprops' in MATLAB computes the perimeter by calculating the distance between each adjoining pair of pixels around the border of the region.
Solidity: The proportion of the pixels in the convex hull that are also in the region. It is computed as Area/Convex Area.
Orientation: The angle (in degrees ranging from -90 to 90 degrees) between the x-axis and the major axis of the ellipse that has the same second-moments as the region.
Extent: Extent specifies the ratio of pixels in the region to pixels in the total bounding box. It is computed as the Area divided by the area of the bounding box.
EulerNumber: It specifies the number of objects in the region minus the number of holes in those objects.
EquivDiameter: It is the diameter of a circle with the same area as the region. It is Computed as sqrt (4*Area/pi).

IV. PROPOSED CLASSIFIER

A. Multilayer Feed Forward Neural Network

An Artificial Neural Network [11] is a computational model inspired in the functioning of the human brain. It is composed by a set of artificial neurons i.e. processing units that are interconnected with other neurons. Every system is basically a 3 layered system, which are Input layer, Hidden Layer and Output Layer. The input layer has neurons which transfer data via synapses to the hidden layer. Similarly the hidden layer transfers this data to the output layer. Multilayer Feed-forward ANNs [8][14] allow signals to travel one way only; from input to output. There is no feedback i.e. the output of any layer does not affect that same layer.

Feed forward networks often have one or more hidden layers of tansig/purelin neurons followed by an output layer of linear neurons as shown in Figure 3. Multiple layers of neurons with nonlinear transfer functions allow the network to learn nonlinear and linear relationships between input and output vectors. Multiple-layer networks are quite powerful. For example, a network of 2 layers, where the first layer is sigmoid and the second layer is linear, can be trained to approx. any function (with a finite number of discontinuities) arbitrarily well.



W= weight matrix, R= number of elements in input vector and S=number of neuron in respective layer
Figure 3. A Multilayer Feed Forward Neural Network model

B. Backpropagation Algorithm

Back-propagation [9] training algorithm when applied to a feed forward multi-layer neural network is known as Back propagation neural network. Functional signals flows in forward direction and error signals propagate in backward direction. That's why it is Error Back Propagation or shortly Back Propagation network. The activation function that can be differentiated (such as sigmoid activation function) is chosen for hidden and output layer computational neurons. Usually a gradient descent algorithm is used to adjust the neural networks weight by comparing the target (desired) and actual network results when a set of inputs are introduced in the network.

C. Network Design

The neural network is designed with 4 layers. The first layer is the input layer, which consists of 12 neurons. Both 1st and 2nd hidden layer consisting 90 neurons each. The output layer consists of a single output neuron specifying the output class of a particular plant. Transfer functions in each layer are purelin, tansig, tansig and logsig respectively which are inbuilt in MATLAB. For training and learning we defined the functions with Backpropagation algorithm. We use mean square error value for performance function with 10000 no. of epochs. The input layer takes the 12 features as input: one feature for each input neuron. These features propagate into the hidden layers and finally to the output neuron, which gives the output class.

D. Training and Testing of Network

A NN is trained to map a set of input data by iterative adjusting the weights iteratively. Optimization of the weights is usually made by backward propagation of the error during training or learning phase [10]. The error in prediction is minimized across many training cycles (epochs) until network reaches specified level of accuracy. Before training weights and biases must be initialize. Once the network weights and biases have been initialized, the network is ready for training. We used random numbers around zero to initialize weights and biases in the network. The training needs a set of proper inputs and targets. During training, the weights and biases of the network are iteratively adjusted to minimize the network performance function.

Experiments were performed and the sizes of the training and testing sets were determined by taking into consideration the classification accuracies. The data set was divided into two different data sets – the training data set and the testing data set. The training data set was used to train the neural network, whereas the testing data set was used to verify the accuracy and the effectiveness of the trained model for the classification. First we train the train file database according to the features values as shown in the Figure 4 and after that goes to the separate test file based on the proposed technique thus the matched leaf comes as output as shown in the figure 5.

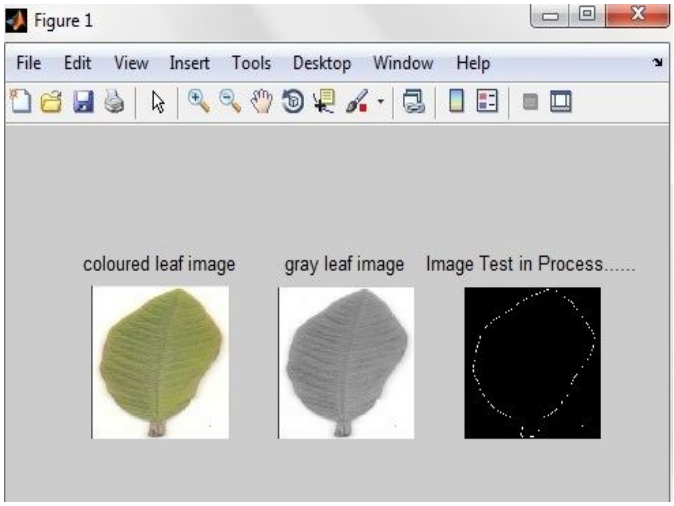


Figure 4. Network Training and testing module for a leaf

As shown in the Figure 6 Training process stops when any one of the following conditions occurs:
1) The maximum number of EPOCHS (repetitions) is reached.
2) The maximum amount of TIME has been exceeded.
3) Performance GOAL has been minimized.

Figure 5. Matched leaf image

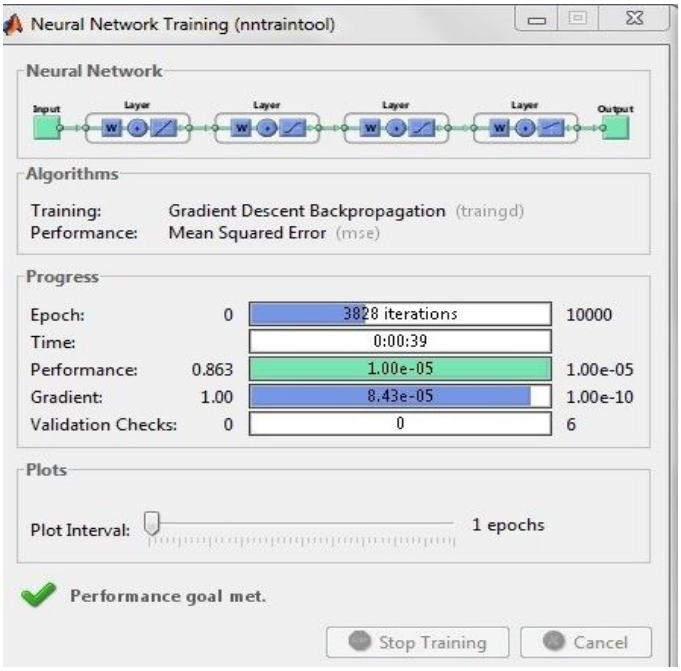


Figure 6. Neural network training and the performance goal met

V. RESULT AND DISCUSSION

All the experiments are programmed by Matlab [18] and run on Intel(R) core i-3 with the clock of 2.40 GHz and the RAM of 4GB under windows 7 environment. The Samples of leaf images belonging to the 16 classes' database is shown in Figure 7. The database contains various leaves with various shapes, colours and size. Experiment was done with these different leaves of different classes and tested in our classifier.

We checked the accuracy that it matches the actual image or not for the database of 440 leaves of different 16 plants species taken, which are completely different in their shape, colour. To each kind of plant 3 leaves from test file are used to test the accuracy. The detailed description of the leaf number of different plant species with their incorrect recognition are listed in Table I. This is used for the calculation of accuracy

The graph between the performance and the number of epochs shows the no. of epochs at which the performance got met as shown in Figure 8 for one leaf matching.

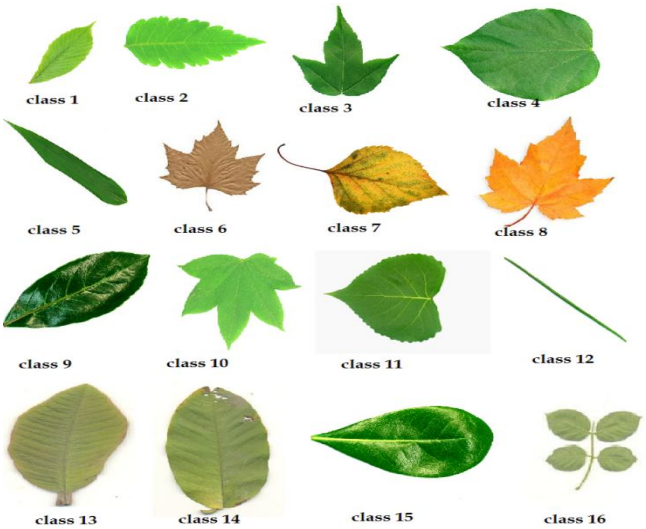


Figure 7. Samples of leaf images belonging to the 16 classes

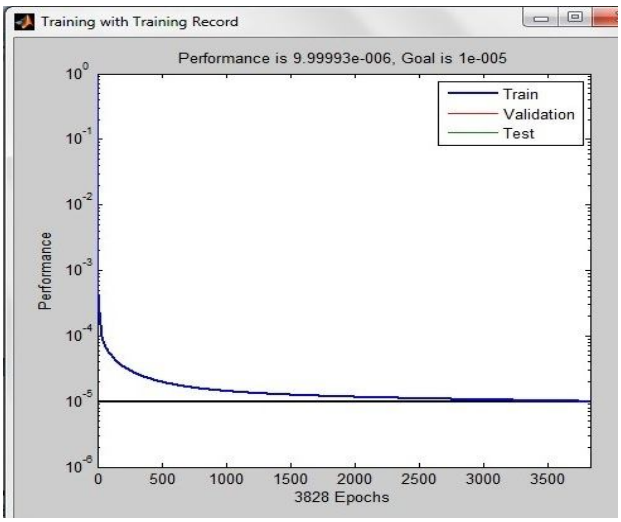


Figure 8. Performance graph Vs no. of epochs for matched leaf

Table 1. Details About the Leaf Numbers of Different Types of Plants

Table 2.

Class	Common Name	No. of leaf samples	No. of incorrect recognition
Class 1	Chestnut leaf	40	2
Class 2	Golden rain tree	47	3
Class 3	Trident maple	53	2
Class 4	Redbud	52	3
Class 5	Bamboo	45	1
Class 6	Eenbruinigherfstblad	23	4
Class 7	Autumn leaf	30	3
Class 8	Golden Maple Leaf	7	5
Class 9	Arrow wood	25	3
Class 10	Castor aralia	30	2
Class 11	Poplar	40	4
Class 12	Deodar	47	0
Class 13	Guava	18	4
Class 14	mango	15	3
Class 15	Castor aralia	38	1
Class 16	Rose	15	3

Out of total sample of 440 leaves of 16 kinds of plants 401 were classified and 39 were misclassified, that is a recognition accuracy of 91.13%.

VI. CONCLUSION

This paper presents an automated recognition system for the plants leaf image using the multilayer feed forward neural network and back propagation algorithm. Experimental result indicates that the technique is workable with accuracy greater than 90% for the 10000 no. of epochs required for training the neural network. This technique is fast in execution, efficient in classification and easy in implementation. As an overall conclusion, this paper is successful as it met the objectives of the paper and successfully developed, run and optimized the performance of the classification technique. In future, we are going to apply statistical pattern recognition methods which take noise into consideration and more plant's database.

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers for their comments and suggestions to improve this paper. We also wish to express our sincere thanks to all faculty members and friends of the Computer Science department for their cooperation and valuable suggestions in each stage, which lead to proceed of our work.

REFERENCES

[1] S. Wu, F. Bao, E. Xu, Y. Wang, Y. Chang, and Q. Xiang, "A Leaf Recognition Algorithm for Plant Classification Using Probabilistic Neural Network," IEEE 7th International Symposium on Signal Processing and Information Technology, December 2007.

[2] Z. Wang, Z. Chi, and D. Feng, "Shape based leaf image retrieval," IEEE Proceedings Vision, Image and Signal Processing, vol. 150, no. 1, February 2003.
[3] H. Fu, Z. Chi, D. Feng, and J. Song, "Machine learning techniques for ontology-based leaf classification," IEEE 2004 8th International Conference on Control, Automation, Robotics and Vision, Kunming, China, 2004.
[4] J. Du, D. Huang, X. Wang, and X. Gu, "Computer-aided plant species identification (CAPSI) based on leaf shape matching technique," Transactions of the Institute of Measurement and Control. 28, 3 (2006) pp. 275-284.
[5] X. Gu, J. Du, and X. Wang, "Leaf recognition based on the combination of wavelet transform and gaussian interpolation," in Proceedings of International Conference on Intelligent Computing 2005, ser. LNCS 3644. Springer, 2005.
[6] X. Wang, J. Du, and G. Zhang, "Recognition of leaf images based on shape features using a hypersphere classifier," in Proceedings of International Conference on Intelligent Computing 2005, ser. LNCS 3644. Springer, 2005.
[7] J. Du, D. Huang, X. Wang, and X. Gu, "Shape recognition based on radial basis probabilistic neural network and application to plant species identification," in Proceedings of 2005 International Symposium of Neural Networks, ser. LNCS 3497. Springer, 2005.
[8] S. Guarnieri, F. Piazza, and A. Uncini, "Multilayer neural networks with adaptive spline-based activation functions," in Proc. of Int. Neural Network Soc. Annu. Meet. WCNN, Washington DC, pp. I695– I699, 1995.
[9] Pai, G. V and Rajasekaran, S, (2006), 'Neural Networks, Fuzzy Logic and Genetic Algorithms Synthesis and Applications', 6th ed, Prentice Hall of India Pvt. Ltd.
[10] L. Vecci, F. Piazza, and A. Uncini, "Learning and approximation capabilities of adaptive spline activation neural networks," Neural Networks, vol. 11, no. 2, pp. 259–270, Mar. 1998.
[11] B. C. Heymans et al., "A neural network for Opuntia leaf-form recognition," IJCNN, vol. 3, pp. 2116-2121, 1991.
[12] Xiao-Feng, Wang, De-Shuang, Huang, Ji-Xiang, Dua,HuanXu, LaurentHutte, "Classification of plant leaves with complicated background", Applied Mathematics and Computation,2008 Vol.205 pp.916–926.
[13] J.-X. Du, X.-F. Wang and G.-J. Zhang, "Leaf shape based plant species recognition," Applied Mathematics and Computation, vol. 185, 2007.
[14] Daniel Svozil, V. KvasniE, J. Pospichal "Introduction to multi-layer feed-forward neural networks", ELSEVIER Chemometrics and Intelligent Laboratory Systems 39 (1997) 43-62.
[15] M. T. Hagan, H. B. Demut, and M. H. Beale, Neural Network Design, 2002.
[16] Anant Bhardwaj, Manpreet Kaur "A REVIEW ON PLANT RECOGNITION AND CLASSIFICATION TECHNIQUES USING LEAF IMAGES", Accepted for publication in International Journal of Engineering Trends and Technology-Volume 4, Issue 2- 2013, pp 86-91.
[17] Anant Bhardwaj, Manpreet Kaur, and Anupam Kumar "Recognition of plants by Leaf Image using Moment Invariant and Texture Analysis" Communicated to International Journal of Innovation and Applied Science, ISSN 2028-9324 Vol. 3 No. 1 May 2013, pp. 237-248.
[18] www.mathwork.com
[19] Jan Flusser, "Rotation Moment Invariants for Recognition of Symmetric Objects," IEEE Trans. On image processing, vol. 15, no. 12, pp. 3784-3790, December 2006.
[20] J. Hossain and M.A. Amin, "Leaf shape identification based plant biometrics", 13th International Conference on Computer and Information Technology (ICCIT), Pp. 458-463, 2010.
[21] S. Prasad, P. Kumar and R.C. Tripathi, "Plant leaf species identification using Curvelet transform", 2nd International Conference on Computer and Communication Technology (ICCCT), Pp. 646–652, 2011.
[22] Abdul Kadir, L.E. Nugroho, A.susanto, P.Insap Santosa, "Experiments of Zernike Moments for Leaf Identification", Journal of Theoretical and Applied Information Technology, 15 July 2012. Vol. 41 No.1.
[23] V. Cheung and K. Cannons, "An Introduction to Probabilistic Neural Networks", http://www.psi.toronto.edu/~vincent/research/presentations/PNN.pdf, (2003).
[24] J. Slens, "A Tutorial on Principal Component Analysis", http://www.snl.salk.edu/~shlens/pca.pdf, (2009).
[25] A. Kadir, L.E. Nugroho, A. Susanto and P. I. Santosa, "Performance Improvement of Leaf Identification System Using Principal Component Analysis" International Journal of Advanced Science and Technology Vol. 44, July, 2012.
[26] M. Z. Rashad, B.S.el-Desouky, and Manal S .Khawasik "Plants Images Classification Based on Textural Features using Combined Classifier" International Journal of Computer Science & Information Technology, Vol 3, No 4. August 2011.

