A Novel Leaf Classification Technique using GLCM and RBFNN

[1]P.Jona Innisai Rani, [2]C.Suresh Kumar, [3]S.Venkatakrishnan
[1]Research & Development Centre, Bharathiar University, Coimbatore
[2]Principal, J.K.K. Nataraja College of Engg and Tech, Komarapalayam
[3]Assistant Professor, Engineering Wing , Annamalai University , Chidambaram

Abstract:-Identification of leaf and plants is an area of research which has gained a lot of attention in these years and is also an important tool in the field of agriculture, crop rotation, cultivation, forestry and much more. The process generally begins with the acquisition of images i.e., enhancement of leaf images, segmentation of leaf, its feature extraction and the classification. Today, Classification of plants using its various categories has been a broad application. In this paper we presentdifferent techniques which can be used for plant leaves classification. The classification method includes some segmentation algorithms and pattern classification techniques. This technique helps in plant-leaf classification. This process and analysis is effective and the performance of the leaf classification system is analyzed using Radial Basis Function Neural Network (RBFNN). RBFNN enables non linear transformation followed by linear transformation to achieve a higher dimension in hidden space. RBFNN is trained and tested for various categories of leaf images using different Grey Level Co-Occurrence Matrix(GLCM) Features. The results show satisfactory performance and the highest accuracy of 93.04% is achieved using Gaussian Kernels.

Keywords : Acquisition; Segmentation Algorithms; GLCM Features; RBFNN;

I. INTRODUCTION

Plants are omnipresent and it is the backbone of all life on Earth and a significant resource for human well-being. The first and foremost step during design phase is leaf recognition which further continued to get the final identification of plant. Plant Identification plays crucial role in various fields like medicine, agriculture, forestry and pharmacological science etc., Due to various serious issues like global warming and lack of awareness of plant knowledge, the leaf categories are becoming rare and many of them are about to extinct. Development of a quick and efficient classification method has become an area of active research. With the current technologies, plants are very well used in the sector of medicine and agriculture. The major challenge in this study is to identify the most valuable and favorable algorithm and techniques for plant identification through leaf recognition.

II. LITERATURE REVIEW

Many methodologies have been developed in automated fashion for plant leaf identification. Image processing techniques performs image segmentation task which identifies the texture, color and pattern of the leaf.Classification of plantswas based on the characterization of texture properties. They have utilized a combined classifier learning vector quantization along with the radial basis function. There also exists a method that incorporates shape and vein features. In Stephen Gang Wuet al (2007) which uses Artificial Neural Network (ANN) for classification. In Liu Huang(2011), recognition feature techniques such as Scale Invariant Feature Transform(SIFT) and Bag of Features(BoF) are considered, which also generates lower dimensional feature vectors. In Abdul Kadir et al(2012) features were extracted using Principal Component Analysis(PCA) and converted into orthogonal features and the results were inputted to the classifier which used Probabilistic Neural Network(PNN). The Shape features were eccentricity, roundness, dispersion, solidity, convexity, and features called Generic Fourier Descriptors (GFDs). In VishakhaMetre et al(2013) uses Stochastic Gradient Descent (SGD) algorithm for segmentation and k- Nearest Neighbour(k-NN) for classifying the plants. In Mohamed ElhadiRahmani et al (2015), comparing the performance of various classification algorithms also considering the supervised learning algorithms for classification especially decision tree and Naïve Bayes.

III. OUTLINE OF THE WORK

The database which contains leaf images are first acquired. There are three major phases which are Preprocessing, Feature Extraction and Classification. The first phase of preprocessing follows a CRE method which is Contrast Improvement, Removal of Noise (Noise Removal) and Edge Enhancement. The preprocessing is extended to segment the input images. Segmentation is carried out using Fuzzy C-Means (FCM) and Discrete Wavelet Transform (DWT). Segmented features are given as input for feature extraction using Grey Level Co-Occurrence Matrix(GLCM). The last phase of Classifying is done using RBFNN(Radial Basis Function Neural Network) technique which identifies the leaf image.

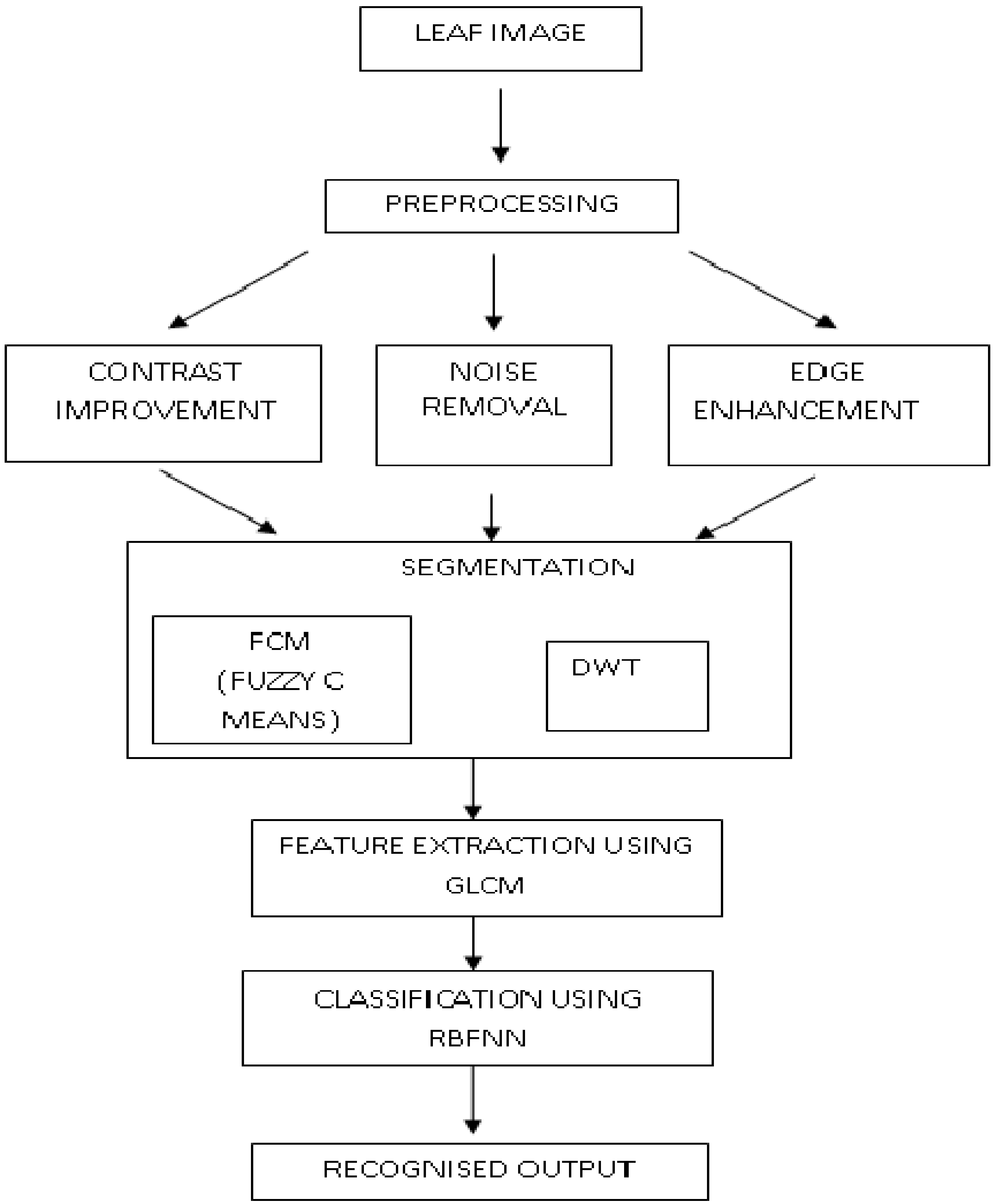

Fig 1. Block Diagram of Leaf Classification

IV. PROPOSED METHODOLOGY

Almost the leaf classification suggests the preprocessing method gives a better performance for the segmentation. The segmentation process goes to using two main methods which is given on the block diagram. Those methods are Fuzzy C-

Means (FCM) and Discrete Wavelet Transform (DWT). This implies the feature extraction inputs. Also some of the features extracted from the input.

V. PREPROCESSING

A. Contrast Enhancement (C)

The first main process of preprocessing is contrast enhancement which is achieved by weighing the input leaf image and the interim equalized image recursively until the allowed intensity range is maximally covered. This technique is known as Histogram Equalization.

B. Removal of Noise (R)

Using a Gaussian filter for noise suppression, the noise is smoothed out, at the same time the signal is also distorted. The use of a Gaussian filter as pre-processing for edge detection will also give rise to edge position displacement, edges vanishing, and phantom edges.

C. Edge Enhancement (E)

Fuzzy C means clustering is a technique which produces high quality images. There are existences of many edge detection methods.Among these this approach using Fuzzy logic elevates the performance in the output for Gray scale images.The features considered for edge detection are Mean, Variance and Correlation. Using these features, edge enhancement is performed.

$$Mean = \mu_i = \sum_{i,j=0}^{N-1} i(P_{i,j}) \tag{1}$$

$$Variance = \sigma_i^2 = \sum_{i,j=0}^{N-1}(i - \mu_i)(P_{i,j}) \tag{2}$$

$$Correlation = \sum_{i,j=0}^{N-1} P\left[\frac{(i-\mu_i)(j-\mu_j)}{\sqrt{(\sigma_i^2)(\sigma_j^2)}}\right] \tag{3}$$

VI. SEGMENTATION

A. Fuzzy C Means Clustering

Fuzzy c-means has been a very important tool for image processing in clustering objects in an image. Mathematicians included the spatial term into the FCM algorithm to improve the accuracy of clustering under noise. A fuzzy logic model can be described on fuzType equation here.zy sets that are defined on three components of the HSL- color space , HSL and HSV. The membership functions aim to describe colors follow the human intuition of color identification.

The Fuzzy C Means Algorithm aims to partition a finite group of n elements X={x_1,…,x_n} into a collection of c fuzzy clusters with respect to certain given criteria. Given a finite set of data, the algorithm returns a list of c cluster centers $C = \{c_1,\ldots,c_c\}$ and a partition matrix $W = w_{i,j} \in [0,1], i = 1,\ldots,n, j = 1,\ldots,c,$ where each element, w_{ij} , tells the degree to which element x_i , belongs to cluster c_j.

The FCM aims to minimize an objective function

$$arg_c min \sum_{i=1}^{n} \sum_{j=1}^{c} w_{ij}^m \, x_i|| - c_j||^2 \tag{4}$$

Where

$$\omega_{ij} = \frac{1}{\sum_{k=1}^{c}\left(\frac{\|x_i - c_j\|}{\|x_i - c_k\|}\right)^{\frac{2}{m-1}}} \tag{5}$$

B. Discrete Wavelet Transform

The most recent and desirable solution to overcome the shortcoming of traditional transforms is the wavelet transform. The wavelet allows us to trade off the time and frequency resolution in different ways.In discrete wavelet transform, long time windows are considered to get a fine low-frequency resolution, and short time windows are found to get high-frequency information. Thus, wavelet transform gives precise frequency information at low frequencies and precise time information at high frequencies. This makes the wavelet transform suitable for the analysis of irregulardata patterns, such as for leaf images taken at varying instances.Low and medium spatial frequencies usually match image content while high-frequency coefficients usually represent noise or texture areas. So, in wavelet domain you have an additional chance to distinguish image content and noise.

VII. FEATURE EXTRACTION

The task of the feature extraction is to obtain the most relevant information from the original data and represent that information in a higher dimensionality space.The goal of feature selection is to reduce the dimensionality of vectors associated to patterns by selecting a subset of attributes smaller than the original.Gray Level Co-occurrence Matrix (GLCM) exploits the higher-order distribution of gray values of pixel that are defined with a specific distance or neighborhood criterion.The GLCM normalizes each value in the matrix by dividing the total number of occurrence, providing the probability of occurrence of a pair of gray values separated by a distance vector.

The GLCM which is generated is used to derive several properties from it. The most commonly used texture based features are as follows.

1. Homogeneity: Returns a value that measures the closeness of the distributionof elements in the GLCM to the GLCM diagonal. Range is [0 1]. Homogeneity is 1 for a diagonal GLCM.

$$f_1 = \sum_{i,j} \frac{p(i,j)}{1+|i-j|}$$

2*. Autocorrelation:* The autocorrelation feature of an image is used to evaluatethe fineness or roughness of the texture present in the image. This function is related to the size of the texture primitive for example the fitness of the texture.

$$f_2 = \sum\sum (i,j)p(i,j)$$

3. Dissimilarity: Dissimilarity and contrast measure the degree of texture smoothness.

$$f_3 = \sum_i \sum_j \left(i+j-\mu_x-\mu_y\right)^3 p(i,j)$$

4. Entropy: Entropy is a statistical measure of randomness that can be used tocharacterize the texture of the input image.

$$f_4 = \sum_i \sum_j |i-j| p(i,j)$$

5. Sum of squares**:** This feature puts relatively high weights on the elementsthat differ from the average value of P (i, j).

$$f_5 = \sum_i \sum_j (i-\mu)^2 p(i,j)$$

6. Sum Average: Sum of average is the sum of all values and divided by the totalnumber of values and is given by

$$f_6 = \sum_{i=2}^{2N_g} i p_{x+y}(i)$$

7. Sum variance: Variance is a measure of the dispersion of the values aroundthe mean and combinations of reference and neighbor pixels. The extractedfeature Sum of variance is given by

$$f_7 = \sum_{i=2}^{2N_g} (1-f_s)^2 p_{x+y}(i)$$

8. Sum Entropy: Sum entropy is calculated as the summation for all the pixelvalues. The extracted feature Sum of variance is given by

$$f_8 = -\sum_{i=2}^{2N_g} i p_{x+y}(i) \log p_{x+y}(i)$$

9. Information measure of correlation 1:

$$f_9 = \frac{HXY - HXY_1}{\max(HX, HY)}$$

10. Information measure of correlation 2:

$$f_{10} = (1-\exp[-2.0(HXY_2 - HXY)])^{1/2}$$

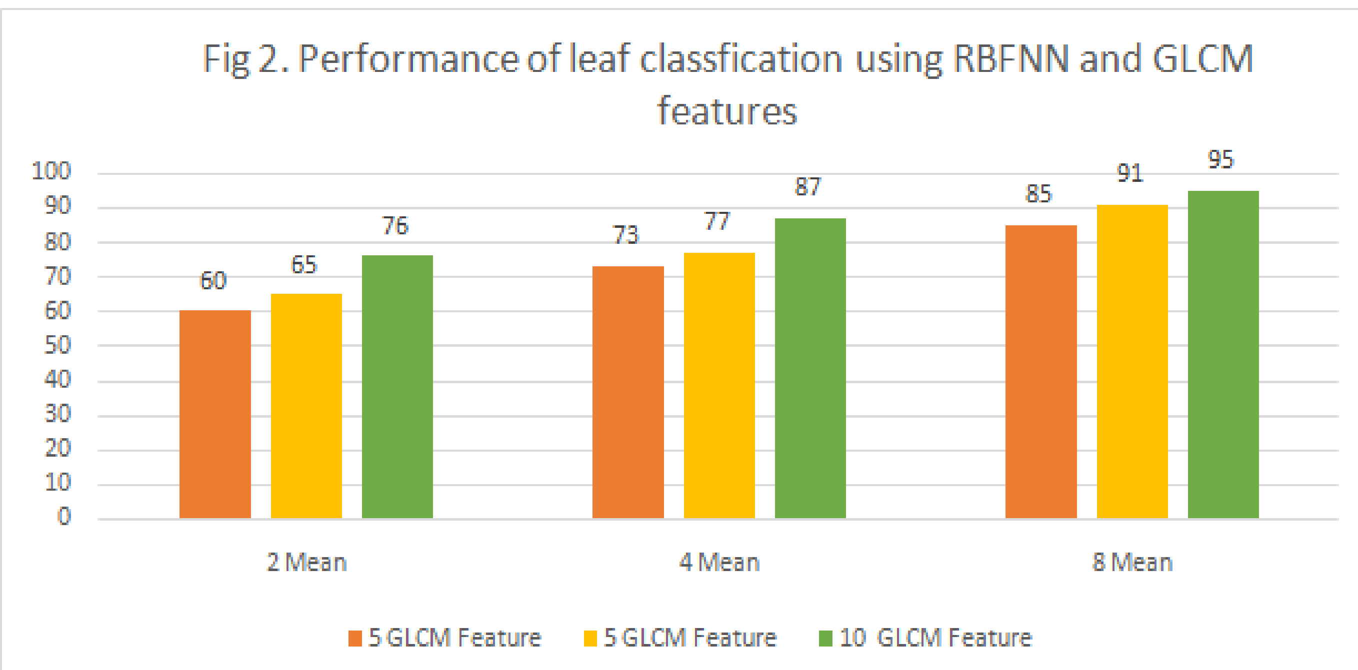

Fig. 2 Performance of leaf Classification using RBFNN and GLCM Features

The performance of leaf classification for varying features are compared and we find that highest performance is obtained when 10 features are taken for consideration.

VIII. CLASSIFICATION

Radial Basis Function Neural Network (RBFNN) possesses the feedforward architecture that consists of an input layer , a hidden layer, and an output layer.A set of inputs and outputs characterize such a network. Hidden units are present in between the inputs and outputs. These hidden units are responsible for implementing the radial basis function. There are N_i units in the input layer for N_i dimensional input vector. N_h hidden layers are connected with these input units and are finally connected to these N_c output layer units, and N_c denotes the total output classes.

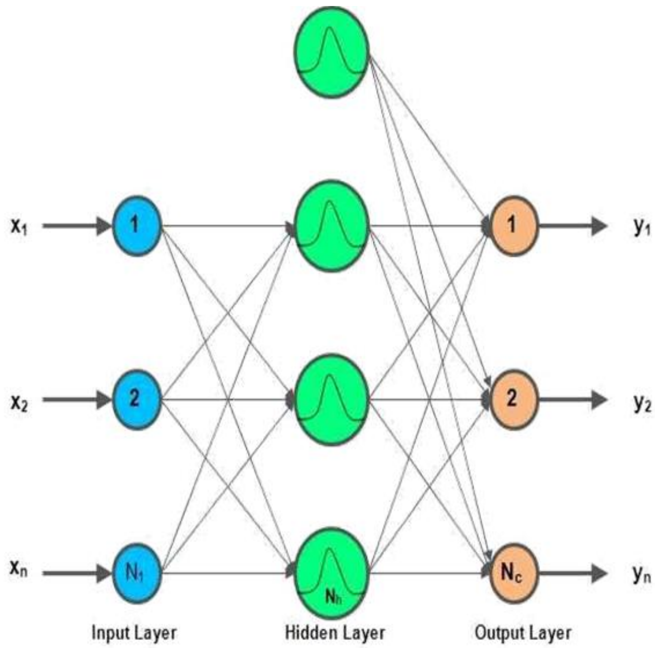

Fig.3 Radial Basis Function Neural Network

IX. EXPERIMENTAL RESULTS

The accuracy in percentage are tabulated. By considering 5, 8 and 10 GLCM features, the accuracy obtained are varying. And we find that 10 mean features performed better when compared to 5 and 8 means. The maximum accurate performance obtained was 95%. The table below shows the accuracy of the kernel functions.

RBFNN	Accuracy in (%)
5 GLCM Features	76
8 GLCM Features	87
10 GLCM Features	95

Table 1: Accuracy of the Kernel Functions.

X. CONCLUSION

In this paper an efficient method for classifying leaf images has been described. 5, 8, 10 dimensional histogram features where extracted from segmented leaf images. The performance of the system was studied using support vector machine. RBFNN was trained and tested for different mean functions and the system showed an accuracy of 95%.

REFERENCES

[1]. Abdul Kadir, Lukito Edi Nugroho, AdhiSusanto and Paulus InsapSantosa, “Performance Improvement of Leaf Identification System Using Principal Component Analysis,” International Journal of Advanced Science and Technology, Vol. 44, pp. 113 - 124, July, 2012.

[2]. VishakhaMetre, JayshreeGhorpade, “An Overview of the Research on Texture Based Plant Leaf Classification,” International Journal of Computer Science and Network, Vol 2, Issue.3, pp. 2277 - 5420, 2013.

[3]. Mohamed ElhadiRahmani, Abdelmalek Amine, Mohamed RedaHamou, “Plant Leaves Classification,” The First International Conference on Big Data, Small Data, Linked Data and Open Data, 2015, pp.75 - 80.

[4]. Suman S. G, B. K. Deshpande, “Plant leaf Classification using Artificial Neural Network Classifier,” International Journal of Innovative Research in Computer and Communication Engineering, Vol. 5, Issue. 5, pp.10189 - 10196, May 2017.

[5]. Ms. DivyaT,Asha C Korwar ,Prof. Virupakshappa, “A Survey On Recognition Of Plant Leaves Using Various Classifiers And Various Morphological Features,” International Journal of Advance Foundation and Research in Computer, Vol. 2, Issue. 5, pp. 89-95, May - 2015.

[6]. Sapna Sharma, Dr. Chitvan Gupta, “Recognition of Plant Species based on leaf images using Multilayer Feed Forward Neural Network,” International Journal of Innovative Research in Advanced Engineering, Vol. 2, Issue. 6, pp. 104 - 110, June 2015.

[7]. Manisha M. Amlekar, Ashok T. Gaikwad, “Plant classification based on leaf Shape features using Neural Network,” International Conference on Innovation in Science, Engineering, Humanities and Management, pp. 214 - 218, September 2017.

[8]. AamodChemburkar, AnandSartape, AjinkyaGawade, Prasad Somawanshi, “Automated Tool for Plant Leaf Classification Using Morphological Features,” International Journal Of Engineering And Computer Science, Volume. 3, Issue. 11, pp. 9098 - 9102, November, 2014.

[9]. A.H. Kulkarni1, Dr. H.M.Rai, Dr. K.A.Jahagirdar, P.S.Upparamani, “A Leaf Recognition Technique for Plant Classification Using RBPNN and Zernike Moments,” International Journal of Advanced Research in Computer and Communication Engineering, Vol. 2, Issue.1, pp. 984 – 988, January 2013.

ISSN No:-2456 –2165

[10]. Stephen Gang Wu, Forrest Sheng Bao, Eric You Xu, Yu-Xuan Wang, Yi-Fan Chang and Qiao-Liang Xiang, “A Leaf Recognition Algorithm for Plant Classification Using Probabilistic Neural Network,” pp. 1 – 6, 29 July 2007.

[11]. A. Kadir,L.E. Nugroho, A. Susanto and P.I. Santosa, “A Comparative Experiment Of Several Shape Methods In Recognizing Plants,” International Journal of Computer Science & Information Technology, Vol. 3, Issue. 3, pp. 256 - 263, June 2011.

[12]. Kue-Bum Lee and Kwang-Seok Hong, “An Implementation of Leaf Recognition System using Leaf Vein and Shape,” International Journal of Bio-Science and Bio-Technology Vol. 5, Issus. 2, pp.57 – 66, April, 2013.

[13]. Abdul Kadir, Lukito Edi Nugroho, AdhiSusanto, Paulus InsapSantosa, “Leaf Classification Using Shape, Color, and Texture Features,” International Journal of Computer Trends and Technology, pp. 225 - 230, Aug 2011.

[14]. AbdolvahabEhsanirad And Sharath Kumar Y. H., “Leaf recognition for plant classification using GLCM and PCA methods,” Oriental Journal of Computer Science & Technology, Vol. 3, Issus 1, pp. 31-36 ,2010.

[15]. JyotismitaChaki, Ranjan Parekh, “Plant Leaf Recognition using Shape based Features and Neural Network classifiers,” International Journal of Advanced Computer Science and Applications, Vol. 2, Issue.10, pp. 41 -47, 2011.

[16]. AnujaVaidya ,DivyeshPujari, Harsh Desai, KrunalBorse and Sagar Patel, “Leaf Recognition – A Technical Review,” International Journal For Research In Emerging Science And Technology, Vol. 2, Issue. 1, pp. 46 – 51, January-2015.

[17]. ShayanHati, Sajeevan G, “Plant Recognition from Leaf Image through Artificial Neural Network,” International Journal of Computer Applications, Vol. 62, Issue.17, pp. 15 – 18, January 2013.

[18]. SubhankarGhosh, Harish Kumar, Pradeep Kumar, Devaraj, JyothiS.Nayak, “Study on Classification of Plants Images Using Combined Classifier,” International Journals of Advance Research in Computer Science and Management Studies, Vol. 3, Issue. 4, pp. 516 – 526, April 2015.

[19]. Mr. GirishAthanikar, Ms. PritiBadar, “Potato Leaf Diseases Detection and Classification System,” International Journal of Computer Science and Mobile Computing, Vol. 5 Issue. 2, pp. 76-88, February- 2016.

[20]. L. Ramesh, V. Nirmala, “Betel Leaf Disease Classification Using Clustering Method,” Journal of Chemical and Pharmaceutical Sciences, Issue. 1, pp. 193 - 196 , February 2017.