RESEARCH ARTICLE OPEN ACCESS

# Enhancing Accuracy of Plant Leaf Classification Techniques

C. S. Sumathi*, A. V. Senthil Kumar **
*(Assistant Professor, School of Post Graduate Studies, Tamil Nadu Agricultural University, Coimbatore-3, Tamil Nadu, India.)
** (Director, Department of Post Graduate and Research in Computer Applications, Hindusthan College of Arts and Science, Coimbatore – 641028, Tamil Nadu, India)

## ABSTRACT
Plants have become an important source of energy, and are a fundamental piece in the puzzle to solve the problem of global warming. Living beings also depend on plants for their food, hence it is of great importance to know about the plants growing around us and to preserve them. Automatic plant leaf classification is widely researched. This paper investigates the efficiency of learning algorithms of MLP for plant leaf classification. Incremental back propagation, Levenberg–Marquardt and batch propagation learning algorithms are investigated. Plant leaf images are examined using three different Multi-Layer Perceptron (MLP) modelling techniques. Back propagation done in batch manner increases the accuracy of plant leaf classification. Results reveal that batch training is faster and more accurate than MLP with incremental training and Levenberg–Marquardt based learning for plant leaf classification. Various levels of semi-batch training used on 9 species of 15 sample each, a total of 135 instances show a roughly linear increase in classification accuracy.
***Keywords:*** *Back Propagation, Incremental back propagation, Levenberg–Marquardt Algorithm, Multi-Layer Perceptron (MLP).*

## I. Introduction

Plant features like fruit, seed, leaf, flower, root, stem etc., help in identifying a plant. As the shape of plant leaves in one of the most important features for characterising various plants visually, the study of leaf image retrieval schemes will be an important stage for developing a plant identification system [10]. The existing electronic herbarium identifies species based on the taxonomic inputs from the user , but it is essential to have a mechanized leaf recognition system, for easy access of the public. There are many ways to plant identification. The conventional methods commonly used are expert determination, recognition, comparison and use of keys and similar devices. These methods are advantageous in their own way [1].

The expert determination is the best option in terms of reliability or accuracy for leaf classification. However, the identification process may consume considerable amount of time even for the experts. Recognition is also considered reliable, next to expert determination. But there are cases where the method becomes inapplicable. Comparison is also consistent but extremely time-consuming. Given a large data set, comparing two plants at a time would be virtually impossible. The option is claimed to be the most widely used method since it does not require much time, materials or experience unlike other methods[1].

An Artificial Neural Network (ANN) is an information processing paradigm that is inspired by the way biological nervous systems, such as the brain, process information. ANN is the computational model formed from several of single units, artificial neurons, connected with coefficients (weights) which constitute the neural structure. The key element of this paradigm is the Processing Elements (PE) as they process information. Each PE has weighted inputs, transfer function and one output. PE is essentially an equation which balances inputs and outputs. It is composed of large number of highly interconnected processing neurons working in union to solve specific problems. ANNs learn by example like human. An ANN is configured for a specific application, such as pattern recognition or data classification, through a learning process. Learning in biological systems involves adjustments to the synaptic connections that exist between the neurones. ANN is a system loosely modelled on the human brain [2].

Many types of neural networks are designed, all can be described by the transfer functions of their neurons, by the training or learning algorithm (rule), and by the connection formula. A single-layer neuron is not able to learn and generalize the complex problems. The MLP overcomes the limitation of the single-layer perceptron by the addition of one or more hidden layer(s).

The MLP has been proven to be a universal approximator [3], a feed forward multilayer perceptron network was presented. MLPs are feed-forward neural networks organised in layers. The input layer consists of distribution points, one or more hidden layers of artificial neurons and one output layer of artificial neurons (nodes). Each node in a layer is connected to all other nodes in the next layer and has a weight. MLPs are often trained using error back propagation. The activation (transfer) function acts on the weighted sum of the neuron's inputs and the most commonly used transfer function is the sigmoid (logistic) function.

## II. Related works

Gradient-based optimization algorithms are the standard methods for adapting the weights of neural networks [4]. The natural gradient gives the steepest descent direction based on a non-Euclidean, from a theoretical point of view more appropriate metric in the weight space. They empirically compared Rprop(resilient back propagations) using the Euclidean and non-Euclidean metric respectively. As their work is closely related to Levenberg-Marquardt learning, this method is added for comparison. Rprop based on the non-Euclidean metric shows at least similar performance as Levenberg-Marquardt learning on the two benchmark problems considered and appears slightly more robust. In both benchmark problems the task of learning the sample data is considered and not the import issue of generalization. It turned out that the Rprop algorithm can indeed profit from using the natural gradient, although the updates done by Rprop are not collinear with the gradient direction. Natural $iRprop^+$ shows similar performance as Levenberg-Marquardt learning on two test problems. The results indicate the natural $iRprop^+$ is little bit faster in the early stages of optimization. Levenberg-Marquardt learning and Rprop using the natural gradient computing a weight update requires cubic time and quadratic space.

A new incremental learning method for pattern recognition called 'Incremental back propagation learning network' was proposed by Limin [5]. An incremental learning system learns y based on x, then learns z based on y and son on. The standard back propagation network is not an incremental learning by its nature. Two problem domain were used to evaluate the learning system. There were 150 instances of three classes viz., setose, verlicolor and viginica, belonging to Iris flowers. Four attributes such as septal length, septal width, petal length and petal width were considered for each instance. Each attribute of all instances was discriticized into three levels. The second domain was the recognition of promoters in DNA nucleotide strings. The performance of an incremental learning system was evaluated in respect of memorization of old knowledge and generalization to unseen instances. The learning curves in all these cases show smooth convergence with minor fluctuation through the process. Thus a new incremental learning method for pattern recognition IBPLN, which employs bounded weight modification and structural adaption learning rules and applies initial knowledge to constraint the learning rule has been proposed in this work.

The analysis of two training algorithm Bayesian Regularization and Levenberg-Marquardt based on MLP neural network reveals that MLP can solve difficult and diverse problem in supervised manner with error back-propagation algorithm [6]. In back propagation algorithm error is to back propagate to adjust the weights to reduce the error between the actual output and estimated output. In their analysis they have simulated the MLP network and computed the localization error. Artificial Neural Network with 3 dimensional inputs have been used and one hidden layer with 15 neurons and two outputs. 121 data sample have been created for training the network. The Bayesian regularisation algorithm is more accurate as compared to Levenberg-Marquardt algorithm. The algorithm also reduces the need for lengthy cross-validation. It has an efficient criterion for stopping the training process and prevents overtraining of the network. This ability makes it a more adaptive and robust back-propagation network for evolving localization algorithm for wireless sensor network. The simulation results demonstrate the effectiveness of the proposed model on localization error.

The effect of two causal factors viz., coating weight gain and amount of pectin-chitosan in the coating solution of the vitro release profile of theophylline for biomodal drug delivery was modelled by incorporating ANN multilayer perceptron feed forward network and developed a predictive model of formulation [7]. Five different training algorithms of three classes, gradient descent, Levenberg-Marquardt and Genetic Algorithm (GA) were used to train NN containing a single hidden layer of four nodes. Subsequently, the performance of aforementioned algorithm was compared with regard to predicting ability. The ANNs were trained with those algorithms using the existing experimental data as the training set. Though GA is often useful in a robust evaluation of the best region of a solution space; it is inefficient and ineffective in fine-tuning local search with their problem's region. Further, training by GA often requires relatively long computational time. Incremental backpropagation and batch backpropagation outperformed the others. Gradient descent backpropagation algorithm in particular incremental and batch propagation can be used as the training algorithms for modelling and prediction of in vitro drug release profiles.

The concept of supervised learning in multi-layer perceptrons based on the technique of gradient descent was introduced by Wilson and Martin [8]. Few problems and drawbacks of the original backpropagation learning procedure are discussed and more sophisticated technique is developed. The performance of several algorithms is tested in twenty runs with different initial weight setting. The fast and robust convergence of adaptive learning algorithms, and the failure of pure gradient, demonstrates the ability of advanced techniques to solve very complex learning tasks. Thus this article gives an overview over past and recent developments in algorithms for supervised learning in multi-layer perceptron.

The use of entropy as a cost function in the neural network learning phase usually implies that, in the back-propagation algorithm, the training is done in batch mode [9]. They present a way of combining both modes when using entrophic criteria, taking profit of the advantages of both methods. The batch-sequential algorithm tries to combine the two methods applied in the back propagation learning algorithm, the sequential mode and the batch mode where the update is performed after the presentation of all samples of the training set. The experiments show that this is a valid approach that can be used to speed-up the training phases, maintaining a good performance.

An application was developed using Cany Edge Detection and multi-layer perceptron for recognizing leaves of topical plants [10]. It recognizes a plant from an input image file using the plant leaf's shape. Hybrid modelling techniques is used to extract features from the leaf. The moment-invariant method is used to extract the first four moments of the image while the centroid-radii method is used to extract 36 radii with respect to the images centroid. Canny Edge Detection technique is used in extracting the edges of the leaf images, which undergo pattern recognition process using multi-layer perceptron. It lists the possible matches for the plant species depending on the training set. This application would be helpful to the researchers and botanist.

Artificial neural network is used to identify plant by inputting leaf image is described by Hati [11]. A new input features and image processing approach that matters in efficient classification in artificial neural network have been introduces compared to earlier approaches. Image processing techniques are used to extract leaf shape features such as aspect ratio, width ratio, apex angle, apex ratio, base angle, centroid deviation ratio, apex ratio and circularity. These extracted features are given as input to neural network. 534 leaves of 20 species of plants were collected, out of which 400 leaves were trained and 234 testing samples were recognized with 92% accuracy.

# III. Material and methods

## 3.1 Feature Extraction

Edge detection is the process of detecting the pixels in the image that represent the edges of the image object. This edge detection process consists of three steps such as: filtering, enhancement and detection. Noise in the image removed during filtering due to random variation in intensity value. Further improvement intensifies the pixels while there is a change in local intensity. Edges are detected using thresholding concept. Prewitt edge detection, Robert edge detection, Sobel edge detection and canny edge detection are most commonly used detection methods and Sobel edge detector is proposed here.

The Sobel edge detector finds the approximate absolute gradient magnitude to detect edges at each point. Regions of high spatial frequency corresponding to edge are obtained by the 2-D gradient measurement. A series of gradient magnitudes can be created using a simple convolution kernel and this convolution can be mathematically represented as,

$$N(x,y)=\sum_{k=-1}^{1}\sum_{j=-1}^{1}K(j,k)\,p(x-j,y-k) \quad \text{—} \qquad (1)$$

The Sobel detector uses two convolution kernels for detecting changes in horizontal contrast (hy) and vertical contrast (hx).

Gabor filters are bandpass filters which are used in image processing for feature extraction, texture analysis. Its impulse response is defined by a harmonic function multiplied by a Gaussian function. Thus, a bidimensional Gabor filter constitutes a complex sinusoidal plane of particular frequency and orientation modulated by a Gaussian envelope [12]. It achieves an optimal resolution in both spatial and frequency domains. The impulse response of these filters is created by multiplying a Gaussian envelope function with a complex oscillation.

Let $x=[x1x2]T$ be the image coordinates. The impulse response of a Gabor filter g(x) is then given by,

$$g_{mn}(x)=\frac{1}{2\pi a_n b_n}e^{-\frac{1}{2}x^T A_{mn}x}\,e^{jK_0{}^T{}_{mn}x} \quad \text{—} \qquad (2)$$

A Gabor function is a sinusoidal modulated Gaussian in the spatial domain. For a 2-D Gaussian curve with a spread of $\sigma_x$ and $\sigma_y$ in the x and y directions, respectively, and a modulating frequency of $u_0$, the real impulse response of the filter [15] is given by,

$$h(x,|\,y)=\frac{1}{2\pi\sigma_x\sigma_y}\exp\left\{-\frac{1}{2}\left[\frac{x^2}{\sigma_x^2}+\frac{y^2}{\sigma_y^2}\right]\right\}.\cos(2\pi u_0 x) \quad \text{—} \qquad (3)$$

**3.2 Feature Selection**

Feature selection is the process of deciding on a subset of relevant features for use in system construction. The feature selection technique is deployed because the data contains many redundant or irrelevant features in it. Redundant features are those which provide only information about the currently selected features and irrelevant features provide no constructive information. It is important to indulge feature selection as it improves the performance of classification algorithm and allows understanding the domain better. Given a set of features V and a target variable T, the minimum set F that achieves maximum classification performance of T gives the best feature.

Correlation based Feature Selection (CFS) is a filter algorithm which ranks feature subsets based on a correlation based heuristic evaluation function. The bias of the evaluation function is toward subset which includes a feature that is highly correlated with the class and uncorrelated with each other. Empirical evidence from the feature selection literature shows that, along with irrelevant features, redundant information.CFS algorithm that couples feature evaluation formula with an appropriate correlation measure and a heuristic search strategy[13].

A feature $V_i$ is said to be relevant iff there exists some $v_i$ and c for which

$p(V_i = v_i) > 0$ such that

$$p(C = c \mid V_i = v_i) \neq p(C = c) \quad — \quad (4)$$

**3.3 MLP with various learning algorithms**

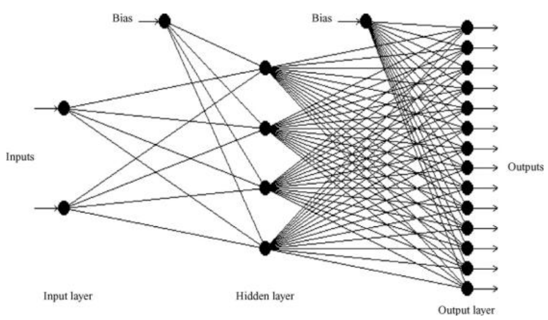


Figure 1. Schematic representation of a multilayer perceptron feed forward network consisting of two inputs, one hidden layer with four neurons and 14 outputs

Fig. 1 shows MLP feed forward network. The network consists of two inputs, one hidden layer with four neurons and 14 outputs.In MLP, various learning algorithms are available such as Back Propagation Learning, Quasi-Newton method [14]. ANN learning paradigms can be classified into supervised, unsupervised and reinforcement learning. Supervised learning model is the availability of a supervisor who classifies the training examples into classes and utilizes the information on the class membership of each training instance. Some methods are Error correction learning rule and memory based learning rule. Unsupervised learning model identifies the pattern class information. Q-Learning is the example of those methods. Reinforcement learning learns through trial and error interactions with its environment (reward/penalty assignment). Competitive learning rule and Hebbian learning rule is the example of this type of learning [15].

**3.3.1 MLP with Levenberg–Marquardt based learning**

Levenberg-Marquardt Algorithm (LMA) is used to solve non-linear least squared problems. LMA provides a numerical solution to the problem of minimizing a function. Back propagation algorithm utilizes the Levenberg-Marquardt algorithm (trainlm) [16] for training of the network. The trainlm'is a network training function that updates weight and bias values according to Levenberg-Marquardt optimization. The Levenberg-Marquardt consists basically in solving $(H + \lambda I)\ \delta = g$ with different $\lambda$ values until the sum of squared error decreases. So, each learning iteration (epoch) will consist of the following basic steps:

1. Compute the Jacobian (by using finite differences or the chain rule)
2. Compute the error gradient $g = J^tE$
3. Approximate the Hessian using the cross product Jacobian $H = J^tJ$
4. Solve $(H + \lambda I)\delta = g$ to find $\delta$
5. Update the network weights w using $\delta$
6. Recalculate the sum of squared errors using the updated weights
7. If the sum of squared errors has not decreased, discard the new weights, increase $\lambda$ using v and go to step 4.
8. Else decrease $\lambda$ using v and stop.

Variations of the algorithm may include different values for v, one for decreasing $\lambda$ and other for increasing it.

Some advantages of LMA are

1. The learning capability of the LMA is reported to be superior and
2. LMA has rapid convergence advantages [17].
3. The LMA is suitable for medium size datasets and fastest among the other training algorithms.

**3.3.2 MLP with Batch Backpropagation algorithm based learning**

In batch back propagation learning, the accumulated weights change, points in the direction of the true error gradient. This means that a sufficiently small step in that direction will reduce the error on the training set as a whole, unless a local minimum has already been reached. In contrast, each weight change made during continuous training will reduce the error for that particular instance, but can decrease or increase the error on the training set as a whole. Hence, batch training produces better accuracy.

Batch learning proceeds as follows:

1. Initialize all weights to small random values.
2. Repeat
3. For each training example do
4. Forward propagate the input features of the example to determine the MLP'soutputs.
5. Back propagate the error to generate Δwij for all weights wij.
6. End for
7. Update the weights based on the accumulated values Δwij.
8. Until stopping criteria reached.

**3.3.3 MLP with Incremental Backpropagation algorithm based learning**

In incremental approach, the weights are changed immediately after a training pattern. The mode of backpropagation algorithm, after each training example is presented to the system and propagated through it, and error is calculated and all connections are modified in backward manner.

Incremental learning proceeds as follows:
Initialize the weights.
Repeat the following steps.
Process one training case.
Update the weights.

## IV. Experimental results

Nine species of plant leaves were selected [10] with 15 samples for each plant species. Sample image of the plant leaves used is shown in Fig. 2.

Figure 2. Leaf samples used in this work.

Matlab was used to extract the features. The features extracted were used to train the classification algorithms. The features were classified using Incremental Backpropagation, Batch Backpropagation and Levenberg–Marquardt algorithms of Multilayer perceptron. Results reveal that batch training is faster and more accurate than MLP with incremental training and Levenberg–Marquardt based learning for plant leaf classification. Finally, the two most well known coefficients to determine the retrieval efficiency, precision and recall, is calculated as follows:

$$\text{Recall} = \frac{\textit{The total number of relavant images in database}}{\textit{Number of relavant image retrieved}}$$

$$\text{Precision} = \frac{\textit{The total number of image retrieved}}{\textit{Number of relavant image retrieved}}$$

$$\text{f Measure} = 2\,\frac{\textit{Precision} * \textit{Recall}}{\textit{Precision} + \textit{Recall}}$$

The classification accuracy obtained is given in Table 1 and Figure 3. Figure 4 and Table 2 tabulates the precision, recall and f Measure for various algorithms and compared.

Table 1. Classification Accuracy

| Technique Used | Classification accuracy |
|---|---|
| MLP with Levenberg-Marquardt based learning | 90.37% |
| MLP with incremental Backpropagation | 91.85% |
| MLP with Batch Backpropagation | 93.33% |

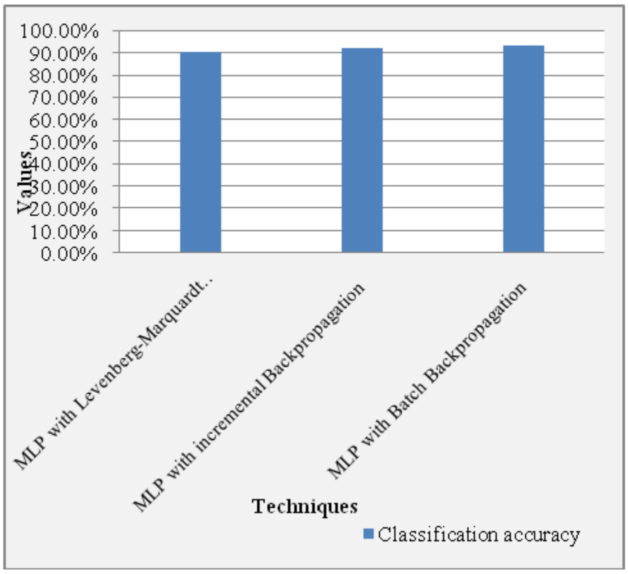


Figure 3. Classification Accuracy

From Table 1 and Fig. 3 it is observed that the classification accuracy is compared for different techniques used with MLP. The classification accuracy of Batch Backpropagation increases by 1.61% than incremental Backpropagation and increases by 3.28% than Levenberg-Marquardt based learning method.

Table 2. Precision, Recall and F Measure

| Technique Used | Precision | Recall | f Measure |
|---|---|---|---|
| MLP with Levenberg-Marquardt based learning | 0.9050 | 0.9037 | 0.9028 |
| MLP with incremental Backpropagation | 0.9225 | 0.9185 | 0.9180 |
| MLP with Batch Backpropagation | 0.9356 | 0.9333 | 0.9335 |

From Table 2 and Fig. 4 it is observed that the precision, recall and Fmeasure values are compared for different techniques used with MLP. The precision of Batch Backpropagation increases by 1.42% than incremental Backpropagation and increases by 3.38% than Levenberg-Marquardt based learning method. The recall of Batch Backpropagation increases by 1.61% than incremental Backpropagation and increases by 3.28% than Levenberg-Marquardt based learning method. The Fmeasure of Batch Backpropagation increases by 1.69% than incremental Back propagation and increases by 3.4% than Levenberg-Marquardt based learning method.

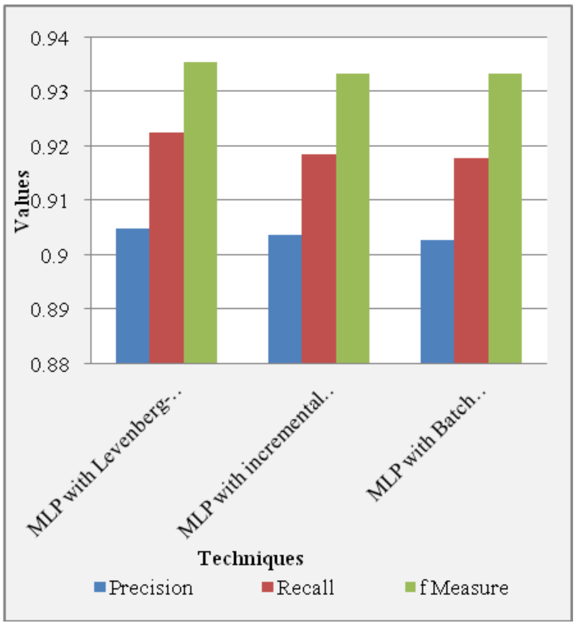


Figure 4. Precision, Recall and FMeasure

## V. Conclusion

In this study, MLP batch propagation was proposed. MLP with incremental back propagation and Levenberg–Marquardt based learning were used in existing method and was compared with the proposed method for estimating the accuracy values. Feature extraction and selection was performed. Nine species of plant leaves were selected with 15 samples for each plant species. From experimental results it is observed that the classification accuracy, precision, recall and Fmeasures were compared for different techniques used with MLP. The classification accuracy of Batch Back propagation increases by 1.61% than incremental Back propagation and increases by 3.28% than Levenberg-Marquardt based learning method.

## References

[1] "Plant identification", [online] available http://www.ibiblio.org/unc-biology/herbarium courses/CHPT 25.html.

[2] Pranob k Charles, Dr. H.Khan, Ch. Rajesh kumar, N.Nikhita.Santhosh , V. Harish, H.Swathi, Artifickal Neural Network based image compression using Leveberg Marquardt algorithm, International Journal of modern Enggineering Research (IJMER), vol 1 issue 2 pp-482-489

[3] Randall Wilson, D., & Martinez, T. R. (2000). The inefficiency of batch training for large training sets. In *Neural Networks, 2000. IJCNN 2000, Proceedings of the IEEE-INNS-ENNS International Joint Conference on* (Vol. 2, pp. 113-117). IEEE.

[4] Igel, C., Toussaint, M., &Weishui, W. (2005). Rprop using the natural gradient. In *Trends and applications in constructive approximation* (pp. 259-272). Birkhäuser Basel.

[5] Fu, L., Hsu, H. H., & Principe, J. C. (1996). Incremental backpropagation learning networks. *Neural Networks, IEEE Transactions on*, *7*(3), 757-761.

[6] Payal, A., Rai, C. S., & Reddy, B. V. R. (2013, January). Comparative analysis of Bayesian regularization and Levenberg-Marquardt training algorithm for localization in wireless sensor network. In *Advanced Communication Technology (ICACT), 2013 15th International Conference on* (pp. 191-194). IEEE.

[7] Ghaffari, A., Abdollahi, H., Khoshayand, M. R., Bozchalooi, I. S., Dadgar, A., &Rafiee-Tehrani, M. (2006). Performance comparison of neural network training algorithms in modeling of bimodal drug delivery. *International journal of pharmaceutics*, *327*(1), 126-138.

[8] Riedmiller, M. (1994). Advanced supervised learning in multi-layer perceptrons—from backpropagation to adaptive learning algorithms. *Computer Standards & Interfaces*, *16*(3), 265-278.

[9] Jorge M. Santos, Joaquim Marques de Sa and Luis A. Alexandre, batch-sequentil algorithm for Neural Networks trained with entropic acrteria, ICANN 2005, LNCS 3697, pp 91-96
[10] SOSA, J. P. M., SOLANO, G. A., & OBICO, J. A. An Application Using Canny Edge Detection and Multilayer Perceptron for Recognizing Leaves of Tropical Plants.
[11] Hati, S. (2013). Plant Recognition from Leaf Image through Artificial Neural Network. *International Journal of Computer Applications*, *62*(17), 15-18.
[12] J. R. MOVELLAN, Tutorial on Gabor filters, http://mplab.ucsd.edu/tutorials/gabor.pdf.
[13] Hall, M. A. (1999). *Correlation-based feature selection for machine learning* (Doctoral dissertation, The University of Waikato).
[14] Ghosh, A., &Chakraborty, M. (2012). Hybrid Optimized Back propagation Learning Algorithm for Multi-layer Perceptron. *International Journal of Computer Applications*, *60*(13).
[15] Sathya, R., & Abraham, A. (2013). Comparison of Supervised and Unsupervised Learning Algorithms for Pattern Classification. *International Journal*.
[16] Tang, L., Tian, L. F., & Steward, B. L. (2003). Classification of broadleaf and grass weeds using Gabor wavelets and an artificial neural network.*Transactions of the ASAE*, *46*(4), 1247.
[17] Anwar, S., Smith, L., & Smith, M. (2012). Innovative machine vision technique for 2D/3D complex and irregular surfaces modelling. *International Journal of Computer Science Issues*, *9*(5), 113-121.