Indonesian Journal of Electrical Engineering and Computer Science
Vol. 20, No. 3, December 2020, pp. 1278~1288
ISSN: 2502-4752, DOI: 10.11591/ijeecs.v20.i3.pp1278-1288

Detection and classification of various pest attacks and infection on plants using RBPN with GA based PSO algorithm

Kapilya Gangadharan[1], G. Rosline Nesa Kumari[2], D. Dhanasekaran[3], K. Malathi[4]
[1,3,4] Department of CSE, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, India
[2]Department of CSE, Shadan Women's College of Engineering and Technology, India

Article Info

Article history:

Received Feb 12, 2020
Revised May 2, 2020
Accepted Jun 8, 2020

Keywords:

Accuracy
Classification
Efficient
Neural Network
Optimization

ABSTRACT

Machine learning methodologies are commonly used in the field of precession farming. It prospects greatly in the plant safety measure like disease detection and classification of pest attacks. It highly influences the crop production and management. The venture of our system is to produce healthy plantation. The proposed system involves Enhanced Feature Fractal Texture Analysis, Statistical Feature Selection and Machine Learning methodology for classification. Hence more than ever there is a need for such a tool that combines image processing methodologies and the Neural network concepts and that is supported by huge cloud of structured data which makes the diagnosis and classification part much easier and convenient. The proposed system recognizes and classifies the plant taxonomy and the infection based on the selected statistical features. The neural network concept followed in our proposed system is focused on Artificial Neural Network which uses Recursive Back Propagation Neural network to speed up the training process as well as reduce multiclass problem in the network and optimize the weights on hidden layers of the Network using Genetic Algorithm based Particle Swarm Optimization technique. We have used MATLAB to implement the concept and obtained better accuracy in disease/pest detection and proved to be an efficient method.

Corresponding Author:

Kapilya Gangadharan,
Department of Computer Science and Engineering,
Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences,
Saveetha nagar, Thandalam, Chennai, Tamilnadu, 602105, India.
Email: kapilya@gmail.com

1. INTRODUCTION

The scientific study of diseases in the plants which contributes in understanding the plant health is plant pathology. Wherein the study throws light on numerous factors that can probably involve in the plant health status such as pathogen identification, disease cycles, economic impacts and epidemiology. Disease in plant disrupts the ecosystem as it involves in the animal food chain including human beings and also that affects agriculture worldwide. The study involves the mechanism that contributes variety of outcome to the agriculture which is estimated to be 15 to 25% of potential crop production is lost due to plant health in 2016 [1].

As we know, agriculture contributes to the growth of the entire country at present it is among the top two farm producers in the world. Plant health has to be studied and rectified so as to increase the productivity and growth as well as the food supply chain. Hence close monitoring of crops are very essential especially it affect production significantly and crop management. Image processing plays the role of distance detection efficiently which identifies the plant disease or pest with the foremost evident. The symptoms are identified on the leaves or any parts of the plants. Therefore, identification and classification of plants, percentage of pest, Symptoms can give successful diagnosis. Machine learning methodology is the best in classification process. Scope of the research is related to capturing, detecting and classifying the disease or insect attack on the plants using Image processing methodology and optimized machine learning techniques.

Quality of the image is very important in classification that significantly affects the performance of the classifier. Our aim is to enhance the performance of the system even in low scale images. Neural Networks are applied on ample of researches to identify the plant diseases. An automated system is designed to identify and classify the diseased leaves like rust and scab using multiple AI techniques [2]. It uses backpropagation Neural Network for pre-processing the images in the first stage. Modified self-organizing feature map and Support Vector Mechanism for segmenting the image. To classify again they have performed Support Vector Mechanism technique. And system performed pretty well with 97.8% of accuracy. Since the system utilizes multiple techniques it seems to be a complex structure but the outcome is very efficient and productive. A filtering method called Gabor filter used on Support Vector Mechanism to improvise the performance. To detect two types of Mildews ANN is formulated [3]. It provides an independent system that follows LMBP method which comprehend the symptoms by texture analysis and thermal normalization.

Another work is proposed to detect infection on any part of the plant on multiple crop system. Each crop system followed an individual method [4]. On vegetables they have experimented with Local Binary Pattern to extract the features and for classification a Neuro-K-Nearest Neighbor has been proposed that includes Backpropagation Neural Network method to train the infected leaves and KNN to test the input images. They have obtained 84.11% accuracy while testing with ANN and 91.54% on testing with Neuro-KNN. Meanwhile Principal Component Analysis is performed to extract the features on the commercial crops by combining it with Discrete Wavelet Transform. Using PCA will reduce the features extracted by DCT and for classification Probabilistic Neural Network Algorithm and Mahalanobis distance is proposed. It produced accuracy of 83.17% on Mahalanobis distance method and average of 86.48% by using Probabilistic Neural Network Algorithm. The system used SVM classifier to identify multiple diseases like leaf spots, leaf rust, Mildews etc., which extracted texture feature based on SVM and segmentation using K-Means clustering and produced an accuracy of 85.33%.

Rothe PR et al., 2015 developed a disease recognition system using patterns and Feed Forward BPNN technique. The system is trained to recognize diseases like Myrothecium, bacterial blight, and Alternaria on Cotton plants. Hu's Moment to extract features to calculate the central moments where functions like Translation, Scaling, Rotation are performed. Feed Forward BPNN with 7 hidden nodes is used for classification which solves the multiple class problem and maps all the input vectors to that of the output vectors. The system produced an overall accuracy of 85% [5]. A Novel Algorithm is proposed to detect the abnormalities and diseases on retina [6]. Which used Region of Interest method to segment the image and Kirsch’s Template Techniques to extract the features and an ANN system is structured using Firefly algorithm for classification which achieved an accurate detection by identifying the cotton wool spots by grouping the affected leaves.

A method to spot the diseased area on any part of the plant is proposed [7] Which used SIFT and WSFTA fusion method that extracts the required feature and for feature selection process it adapted PCA with KNN classifier algorithm which obtained an average accuracy of 95.9% which resulted in identifying various diseases on different plants and different parts of the plants. A color-based confirmation helps the Neural Network to differentiate between Pests or objects with similar characters like size and shape as the target pathogen using the color [8]. The neural network can be trained to identify specific characteristic shades as the proposed system which uses Root Mean Square derivation to detect the rust in the plants. The hybrid techniques execl better in the performance of the system [9]. Most of the machine learning system suffers from multiclass problem. A sysem is designed to identify multiple diseases using SVM and KNN algorithm and obtained the accuracy of 87.3% and 83.6% [10].

Ghaiwat SN et al., 2015 developed a system to detect Powdery mildew, downy mildew, and leaf miner infection on cotton plants. The system used a dataset of 44 sample images. 6 images are used to train and 36 to test. The sample images are decomposed into multiple sub-images. Multiple attributes are obtained from the sub-images individually. BPNN is used to classify the disease with the obtained features [11]. A Leaf spot disease detection system using Radial Basic Function Neural Network technique using 326 sample images from the plant village dataset has been developed. The extracted color attributes are used to train and test the images by applying the Radial Basis Function Neural Network Technique with 3 hidden nodes and 12 Input nodes. 290 Images are trained and 36 images are tested. The time required for training the dataset is 14.0 sec and for testing, it requires 2.97 sec. When the extracted features were trained and tested using SVM the system produces an accuracy of 95% and RBFNN produced 96% [12]. A system that combines multiple features and extractes the statistical feature that reduces the dimentionality of the extracted features [13]. The statistical features are fed as input to the Artificial Neural Network. The algorithm trains

20 images and tested with 45 images and obtained accuracy of 100% but it just has single class classification. A microscopic image-based feature extraction method is proposed to detect rust disease on lentils [14]. It combines local binary pattern and Brightness Bi-Histogram Equalization for enhancing the image to obtain better result.

Artificial deep neural network (ADNN) concept for disease classification has been proposed, which utilizes Recursive Backpropagation Neural Network in combination which uses gradient descent with more number of hidden layers that provides an efficient error function and the process estimates the gradient of error function with that of the weights on the neural network concept. The weights on ANN is adjusted and optimized using Particle Swarm Optimization which adapts the natural selection method by Genetic Algorithm concept to make the algorithm more efficient and stable. The proposed system overcomes the multiclass and optimization problem in the neural network and it is is designed to detect two types of disease like Cercospora leaf spot on Peanut leaves, Bacterial Blight on Paddy leaves and three types of insect attacks like Boll weevil attack on cotton buds, European corn borer on Cotton leaves and Fall army worm attack on peanut leaves.

2. RESEARCH METHOD

The paper presents an integrated method that combines artificial neural network (ANN) using recursive BPNN and Genetic Algorithm based Partial Swarm Optimization algorithm. ANN has the ability to understand, discover and learn non-linear and complex relationships, so it estimates the complexity of connection between the first node and the output node. It can extrapolate the hidden node's substance by learning the nature of the input. It is very intelligent in predicting the unknown facts. ANN generally will never enforce any conditions to the inputs unlike the other prediction methodologies. ANN has even progressed in developing Deep Neural Network which has explored all the energizing and transformational developments in computer vision.

The Recursive Backpropagation Neural Network is used on ANN, which is encouraged to feed until a fixed merit are accomplished. Once it achieves the target then errors are evaluated by propagating it backwards. It is very simple fast and easy to analyze and program. Since there is no fixed parameters except for the number of inputs it vary flexibly and it does not required to understand the system in prior. Just with the input, output and activation values it predicts the inner or hidden layers. It follows power and chain rule based on that it employs and functions on any number of outputs. Genetic Algorithm based PSO method is utilized in ANN to optimize the weight estimation to produce high computational efficiency which is gradient free and it has the ability to find global optima. It reduces the cost function with the weights. When we implement the threshold on RBPNN, the connections pass the information to an aggregate node based on the threshold. Before applying the classification methods, we follow the general Image processing methods like acquisition of image, Standardization, Image Enhancement, Segmentation and Feature extraction and then the machine learning concept as shown in the system architecture Figure 1. The system is propsed to extract the statistical features to avoid dimentional features, over come multiclass problem and optimaztion problem in the neural network.

Algorithm 1: Proposed Methodology
Step 1: Select input image through Image acquisition and standardize the image
Step 2: The sample is in RGB color model Convert it into L*A*B* image
Step 3: Enhance the sample by removing the noise using Sigma filter and apply CLAHE histogram Equalization method to improve the intensity
Step 4: Segment the preprocessed sample using Hybrid Watershed Segmentation with Extended K-Means Clustering
Step 5: Using Mathematical Morphology compute the Gradient of the image horizontally and Vertically
Step 6: On the Morphological Gradient Image apply Enhanced Fusion Fractal Texture Analysis method to extract the features (Using SIFT and WSFTA algorithm) by performing PCA on the extracted key points.
Step 7: Statistical PCA is used for feature selection by computing the cluster of feature values
Step 8: Perform Proposed Artificial Neural Network on the extracted features with Multilayer Perception (MLP), which is trained using Recursive Backpropagation Neural Network
RBPNN = RBPNN (disease_feature, disease_type)
Step 9: Apply Genetic Algorithm based PSO method for an optimized classification of disease
Step 10: Classify the plant topology, type of disease or insect attack or a normal leaf and estimate the percentage of infected area (Ratio of disease and the entire leaf)
Step 11: Compute accuracy with the maximum number of iterations

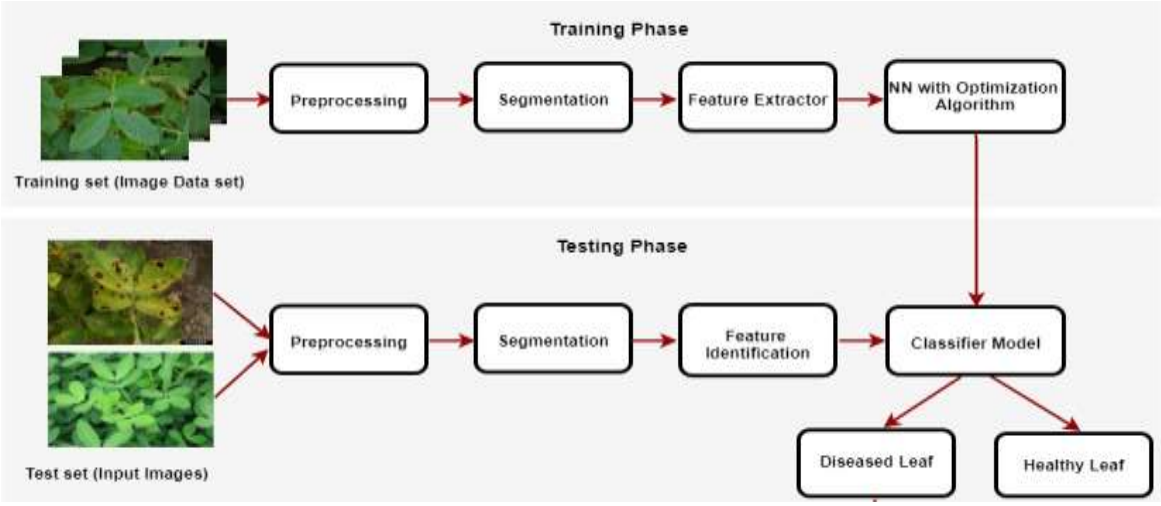

Figure 1. Overall system structure

2.1. Enhanced feature fractal texture analysis and statistical feature selection

Serial based method combines the Key points extracted using SIFT algorithm and Discrete Wavelet Transform based on SFTA. It adopts the inter class and interclass analysis. Entropy is calculated based on the Euclidean Distance, Standard Deviation, Variance, Skewness, Kurtosis, Smoothness these are estimated from the disease affected area. Higher number of features are selected using Principal Component Analysis algorithm by computing the Eigen value and vector for the calculated Entropy by forming a covariance matrix. The below algorithm states the working principal of Principal Component Analysis.

1. Select the value from the MxN matrix
2. Evaluate the correlation or co-variance matrix
3. Estimate the Eigen vector and value of the co-variance matrix
4. Calculate the principal elements and create a feature vector
5. Obtain a new dataset with the vector cluster

Obtained features may not be sufficient to identify the difference between the Interclass and intraclass variance. Hence it is very important to perform feature selection method. Interclass require distribution of lesser number of features for training and intraclass require higher number of features distribution for training. Feature selection needs a non-supervision tool that removes the non-operational features. Because when higher number of features are extracted it might lead to an inappropriate classification. In this paper selection is performed using Statistical-PCA [15]. Steps to perform S-PCA is stated a shown in:

Algorithm 2: Statistical-PCA
Step 1: Choose the nth element from each cluster generated by PCA
Step 2: The extracted features are mapped on the interval (0,1), it sets the features in a specific range
Step 3: Efficient features are selected based on PCA algorithm, it differentiates the nth item of each cluster
Step 4: Repeat step 1 to 3 until N=M (N=1,2,3,4,…M it is number of training data)
Step 5: Pick the best feature that has the most statistical value with reference to step 3

In this process it converts the feature vector of higher dimension to the lower dimension that will limit the occurrence of error.

2.2. Sub section recursive backpropagation neural network

Artificial Neural Network is a registering framework that combines the hardware and the programming language. Its structure depends on the biological neurons' schema and the data processing methodology. The architecture of ANN consists of Artificial Neurons. The Neurons are also known as collection of connected units or nodes. One node can transmit signals to the other node through the intermediate connector. The output of each node is evaluated using non-linear functions. The connectors are also called as edges. The nodes and the edges have weight that corrects the learning process. Depends on the weights on the edges the strength of the connection is evaluated. ANN has three subsets of rules like multiplication, summation and activation. The initial process of ANN is the weight input i.e., each and every generated value should be multiplied with the independent isolated weight and bias using Levenberg- Marquardt BP training method [16]. Then all the calculated weighted values are added using summation and finally the calculated weighted values and bias are passed through activation function. The training is carried out through the Recursive Backpropagation Neural Network. Artificial Neural network that has the recursive topology on the back loops with no limitations until the errors are identified is called Recursive Backpropagation Neural Network. It utilizes internal memory to process the sequence of output recursively. It allows to showcase the dynamic temporal behavior. Backpropagation is a multilayered Neural Network. It is a method that is used for supervised learning [17]. The operation that can be split into two types they are feed forward and Backpropagation

Feed Forward

It is the first phase in ANN [15], where the inputs are transmitted through the number of layers that is hidden and then the final output layer is estimated by vector multiplication on to the hidden layer and the output is multiplied with the input vector of the previous element. Once the input is transmitted from the input to output level through all the hidden nodes present in the network, difference between the target out and the real out is calculated. To estimate the error, we use Recursive Backpropagation Neural Network in our proposed system.

Backpropagation

The Backpropagation is a process of calculating the error in a back loop, in the proposed system recursively we evaluate the error through n iterations until the error is captured it has no limit. The BP training the algorithm is based on the principle of square of Euclidean norms of the output error vector. That is the reason why the NN learning that needs to be optimized by some optimization methods. It reduces the MSE by calculating the required outcome and the hidden layers recurrently or recursively until the missed errors are calculated irrespective of the iterations. We perform faster acquisition methodology. The most essential factor is the learning grade. The value that is chosen should not be either very low or very high an optimal value must be taken always for better accuracy. We multiply momentum to the number of backward iterations. It accelerates the process of learning by changing the weight and continue in the same direction with more steps

Algorithm 3: RBPNN
Step 1: Assign all the input and output
Step 2: Initialize the weights between -1 and 1
Repeat Step 2 for all the available pattern in the training object
Step 3: Implement the input Forward propagation to get the Initial prediction
- i. Check each layer
- ii. Check each and every node in the layer
- iii. Cost implementation is done as below
  - a. Evaluate the sum of weight of input on each node
  - b. Add the threshold to be computed sum
  - c. Calculate the activation for all the available nodes

Step 4: Implement the Backpropagation to propagate the errors through the layers recursively
- i. Check all the output layers
  - a. Estimate the error signal
- ii. Find all the hidden layers and check the available nodes in the hidden layers
  - a. Calculate the error signal for all the nodes
  - b. Update each node with the respective weight
  - c. Repeat above steps recursively until all the error signals are identified and calculated in the hidden layers

Step 5: Calculate the local and global error.
Step 6: Repeat the steps until (maximum number of iterations < specified) and (Error function > specified)
Step 7: Increase the momentum from -0.5 to 0.5 and conduct multiple training sessions.

2.3 Genetic algorithm based partial swarm optimization

Once after the training is completed through RBPNN the weights or the pixels of the input images are considered as the position or the particles in PSO, where we optimize the extracted weights using the RBPNN. A hybrid method that combines Genetic Algorithm and PSO [18] to provide the best optimization. In the first step it initializes the population where the position and the velocity of the particle is considered as the population. The population is extracted from the pixels of the input image. The population can be extracted both locally and globally.

Algorithm 4: GA based PSO
Step 1: Initialize the population (Pixels of the input images are considered as the particle; Population is the position and velocity of the particle) and initialize all GA variables
Step 2: If the Population (PoP) is not improved during N number of iterations then perform the velocity resting process

$$v_i = v_i + (2 \times r - 1) \times v_{max}$$

Where v_i is the vector dimension, i is the particle, r is the uniformly represented random number (0,1) and velocity $v_i$ranges between $[v_{min}, v_{max}]$
Step 3: If $r(0,1) > \gamma$ then create new neighborhood (new neighbor depends upon the percentage of infected region on the image. That is to be optimized) end if
Step 4: Modify the inertia and the particle over the N-Dimensional search space.

$$I\omega = (I_n - I_m) \times \frac{Maximum\ Iteration\ -\ Current\ Iteration}{Maximum\ Iteration} + I_m$$

Where $I\omega$ is the Inertia Weight, $(I_n - I_m)$ are initial and final value of inertia weight
Then Update the information as

i. Evaluate the fitness for each individual x_i
ii. Update the best position Pos_{best} for each individual i
iii. Update the best individual Ind_{best} for each Neighborhood Nh

Step 5: Compute the Velocity and update $V_i(t+1)$ for each individual i and each dimension n

$$v_i(t+1) = \omega v_i(t) + ac_1 r_1 (Pos_{best}(t) - x_i(t) + ac_2 r_2 (Ind_{best}(t) - x_i(t))$$

Where ac_1 and ac_2 are acceleration coefficients, t is time, r_1 and r_2 are the random numbers that are equally distributed with the range (0,1)
Step 6: Compute the current Position $x_i(t+1) = x_i(t) + v_i(t+1)$ for individual i
Step 7: If $r(0,1) > \alpha$ then perform crossover in genetic Algorithm using Euclidian distance and position vector of PSO, else end
Step 8: If $r(0,1) > \beta$ then perform mutation operation else end
Step 9: if GA condition is satisfied with the target condition (number of iteration) then stop the

Where the particles are considered as the birds that flies and lookup for their nourishment in the form of infected part of the image. It searches both globally and locally. Depends on the improvement of the population neighborhood is created and the velocity is rested, that again depends on the percentage of the infected area on the leaf or bud or any other part of the plant. The velocity is calculated using three components, they are current position, variation of the best position (Pos_{best}) and the current position, variation of the best individual (Ind_{best}) and the current position. PSO will speed up the particles towards the best and the accurate local and global solution. Depends on the current position Genetic Algorithm will determine when to use crossover and mutation function [19]. We perform GA operation here to support multi-objective optimization and make the process faster. Also, it works well on noisy substances. It provides approximate solution to optimize and search issues and it is categorized as global search heuristics.

The overall idea of the proposed method is to learn and train the data by adusting the weight and resduce the error as low as possible. The performance of feedforward and back propagation calculates the error and adjust the weight using optimization algorithm so as to match the truth value. The problem with the proposed system is each time the errors are identified; the weights must be multiplied to each hidden neuron. It takes forever in terms of computational time. But it resolves multiclass problem by setting the output based on the number of diseases. It performs better without any misclassification.

3. EXPERIMENTAL RESULTS AND DISCUSSION

In this section the experimental results are computed for the proposed neural network classification.

3.1. Tool

To estimate the efficiency of the proposed system we have tested our algorithm by calculating the elapsed time of execution. The experiment is manifested with the samples that are trained, tested and validated using MATLAB for the proposed system and the existing algorithms. And executed on Windows 10, 64-bit Operating System with 8 GB Ram, 2.70 GHz Intel Core i5 7th Gen Processor.

3.2. Dataset

In this paper we have utilized the datasets available in the internet, the datasets were picked from plant village website [20], Forestry [21] and ICS.UCI.edu [22]. The datasets were passed through all the phase of proposed methodologies mention in this paper. Sample images picked from the dataset is shown in the Figure 2. From the available dataset 80% of the images are used for training, 105 for testing and 10% for validation to obtain better result and efficiency. When 90% of the image is used for training 10% for testing and 10% for validating, it produced an accuracy of 95.26% while the result is improved to 98.76% by performing RBPNN with GA based PSO algorithm that reduced the cost of the generated weight also it has the highest perdition capability. The Artificial Neural Network using RBPNN function employed in the process of diagnosing diseases or insect attack on the plant using a primary dataset as shown in the Table 1.

RBPNN classifier consists of input, hidden and output layers of neurons and it can be learned using the training datasets. The ROC curve demonstrates excellent results on training set which is based on the ROC analysis. The dramatic difference is observed in the testing set which is based on the ratio of images taken for training, testing and validating. From the analysis it is observed that RBPNN classifier with Optimization algorithm produces a better result when compared to the BPNN without Optimization algorithm.

Table 1. Number of Images used in the Dataset

Name of the disease/ Insect attack	Crop Type	Total number of Images	Training	Testing	Validation
Cercospora Leaf spot	Peanut leaves	560	448	56	56
Bacterial Blight disease	Paddy Leaves	540	432	54	54
Boll Weevil	Cotton Bud	480	384	48	48
Fall Army Worm	Peanut leaves	520	416	52	52
European Corn Borer	Cotton leaves and other parts	480	384	48	48
Healthy Leaves	Combined	650	520	65	65

Figure 2. Example of sample images used as dataset from plant village dataset (a) to (e) Cercospora leaf spot on peanut leaves, (f) to (j) Bacterial Blight disease on Paddy leaves, (k) to (0) European corn borer on various parts of cotton plants, (p) to (t) Boll weevil on cotton buds, (u) to (y) Fall arm worm on Peanut leaves

3.3. Experimental Evaluation

Indiscriminate function of the training sets of RBPNN is shown in the Figure 3. Mean square error is estimated from the proposed Neural Network model with different epoch that performs an approximation function. The errors are estimated over the validation process and it is normalized using the number of validation set. Mean square error provides the difference between the actual value and the simulated values. Figure 3 shows the best validation criteria with the number of iterations performed. In our proposed system the best validation performance is 0.50666 at epoch 4. It means after epoch 4 the result would be stable.

The training state shows variation in gradient co-efficient with respect to the number of epochs or number of iterations. The final value of the gradient co-efficient is 0.061239 at epoch 10 which is approximately near zero. Minimum the value of the gradient co-efficient better will be the training and testing criteria. Figure 4 shows the gradient and validation check. It is noticed that as the number of epochs increases the gradient value decreases. When the validation check is equal to 6 at epoch 10, then the number of iterations will be completed.

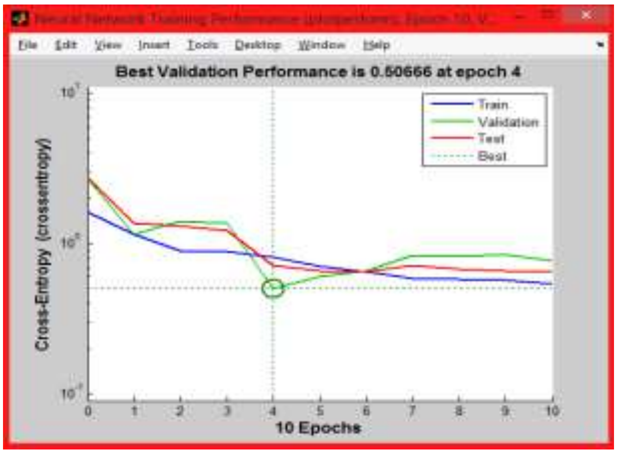

Figure 3. Mean square error plot

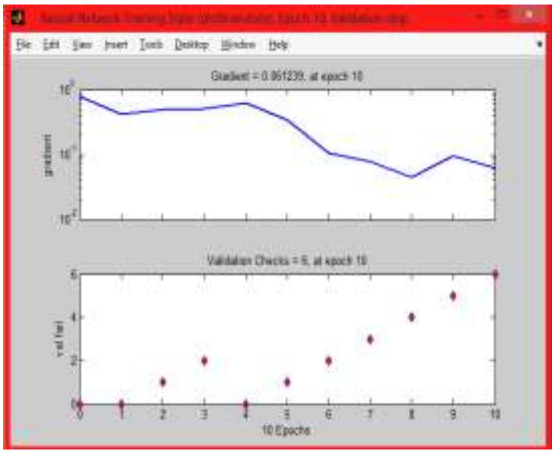

Figure 4. Gradient and validation check

Figure 5 shows the ROC curve for the proposed system performance. ROC was plotted for training, testing and validating the dataset. Where x-axis denotes False positive Rate (0,1), y-axis denotes True positive rate (0,1) and the area is split into two with the landslide slope as shown in the Figure 5 the dotted lines. When the range is between 0 to 0.5 the accuracy will not be calculated properly it produces poor result. The accuracy is perfect and better when the range is maximum and very close to 1 say (0.7 to 1) and it produces the best result. The performance has been evaluated based on the valuation metrics. It can be classified as:

True Positive (TP): Abnormal case appropriately evaluated as Abnormal
False Negative (FN): Abnormal case imperfectly classified as normal
False Positive (FP): Normal case imperfectly identified as abnormal
True Negative (TN): Normal case appropriately calculated as Normal

$$True \text{ Positive } Rate = \frac{True\ Positive}{True\ Positive + False\ Negative} \quad (1)$$

$$False\ Positive\ Rate = \frac{False\ Positive}{False\ Positive + True\ Negative} \quad (2)$$

$$Accuracy = \frac{True\ Positve + True\ Negative}{True\ Positive + True\ Negative + False\ Positive + False\ Negative} \quad (3)$$

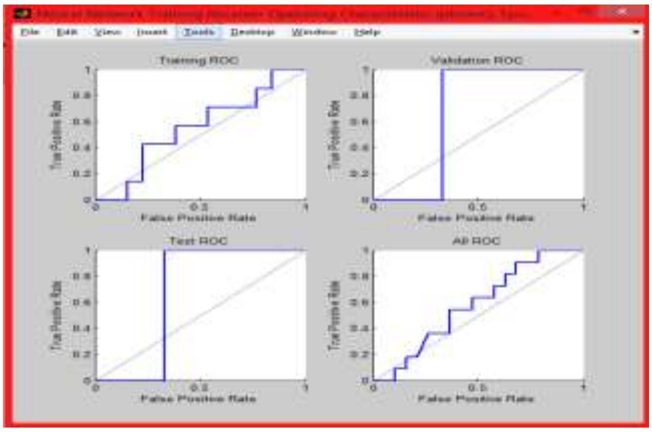

Figure 5. ROC performance evaluation

Figure 6 shows the final classification result for the diseased images which predict the leaf taxonomy, classified the diseased also estimates the affected area so that calculated if the disease is in the early stage, mid stage or the late stage. When the healthy leaves are tested on the same it is identified as No disease found as shown in the Figure 7. When all the types of diseases and healthy leaves were passed through the proposed method it resulted in approximately 98.76% of overall accuracy. Individual accuracy obtained for the five different types of leaves and the healthy leaves are depicted in the Table 2.

Figure 6. Final result of infected leaf

Figure 7. Final result of healthy leaf

Table 2. Result of proposed system

S.No	Name of the disease/ insect attack	% of accuracy
1	Cercospora Leaf spot	100
2	Bacterial Blight disease	98.14
3	Boll Weevil	97.91
4	Fall Army Worm	98.07
5	European Corn Borer	100
6	Healthy Leaves	98.46
	Overall Accuracy	98.76

3.4 Comparitive Evaluation

The proposed system is compared with the existing system and concluded that the proposed system produced better result and performance as shown in the Table 3. For comparison we have chosen multiple classifier algorithms like KNN, SVM, BPNN, RBPNN with optimization algorithm and without optimization algorithm. Tests were conducted for RBPNN with and without optimization algorithm and the other methods were taken from the previous research materials. It is noticed that BPNN performs better when single disease is tested, but when multiple diseases are tested it failed to produce better result. The accuracy of the traditional methods are referred from the existing methods. The comparison between the proposed and the traditional methods are shown in the Table 3.

Table 3. Comparison of proposed algorithm with traditional methods

Classifier algorithm	Disease/insect attack	Number of images		Accuracy
		Training	Testing	
Proposed Algorithm [RBPNN+GA+PSO]	Cercospora, Bacterial Blight, Boll Weevil, fall Army Worm, European Corn Borer	3230	323	98.76
RBPNN without GA+PSO	Cercospora, Bacterial Blight, Boll Weevil, fall Army Worm, European Corn Borer	3230	323	98.45
BPNN [23]	Four Stages of Cercospora	360	40	97.41
SVM [24]	Bacterial, Fungal, Viral, Nematodes and other deficiencies	450	450	93
KNN [25]	Bacterial Blight, Leaf blast and Brown spot	90	30	93.33

4. CONCLUSION

The proposed system resulted in an automatic detection and classification of infections and insect attacks on multiple plants very efficiently. This paper focused mainly on the Statistical Feature selection method and RBPNN algorithm with GA based PSO optimization method. EFFTA extracts the best parameters of texture feature and S-PCA selects the appropriate features that is required for the classifier to distinguish between the healthy and unhealthy leaf patterns. Since the RBPNN is very sensitive to noise we have used sigma noise filter to enhance the image. Image resizing is performed in order to train the dataset from various sources. The optimization problem that occurs in the Neural Network structure s reduced by optimizing the weight along with the constant coefficient of activation function and works on the pattern classification problem. It resolves the variational problems which occurs during the cost estimation that is expressed analytically in terms of unknown functions and multiclass problems. We have trained more than 3000 images and classified the disease and area of infection. Also, our system recognizes the healthy and unhealthy leaves successfully. It performs better with the accuracy of approximately 98.76% with the highest prediction capability. This system works well on the complex background as well.

REFERENCES

[1] Kekane Maruti Arjun, "Indian Agriculture in Indian Economy, Department of Commerce, University of Pune, Pune (Maharashtra)," *India.International Journal of Agriculture and Food Science Technology.*, vol. 4, no. 4, 2013.

[2] Meunkaewjinda A, Kumsawat P, Attakitmongcol K, Srikaew A.," Grape leaf disease detection from color imagery using hybrid intelligent system." *In: IEEE 5th international conference on electrical engineering/electronics, computer, telecommunications and information technology ECTI-CON.*, Krabi, pp 513–516. 2008.

[3] Asefpour Vakilian K, Massah J, "An artificial neural network approach to identify fungal diseases of cucumber (Cucumis sativus L.) plants using digital image processing," *Arch Phytopathol Plant Prot, vol.* 46, vol. 13, pp. 1580-1588, 2013.

[4] Pujari JD, Yakkundimath R, Byadgi AS, " Image processing based detection of fungal diseases in plants." *Procedia Computer Science* , vol. 46, pp. 1802-1808, 2015.

[5] Rothe PR, Kshirsagar RV (2015) Cotton leaf disease identification using pattern recognition techniques. *In: IEEE international conference on pervasive computing (ICPC),* January, pp 1–6.

[6] K. Malathi and R. Nedunchelian, "Detecting And Classifying Diabetic Retinopathy In FundusRetina Images Using *Artificial* Neural Networks-Based Firefly Clustering Algorithm," *ARPN Journal of International Engineering and Applied Sciences*, vol. 11, no. 5, ISSN 1819-6608, March 2016.

[7] Kapilya Gangadharan, G. Rosline Nesa Kumari, D. Dhanasekaran and K. Malathi, "Automatic Detection of Plant Disease and Insect Attack using EFFTA Algorithm" *International Journal of Advanced Computer Science and Applications (IJACSA)*, vol. 11, no. 2, 2020. http://dx.doi.org/10.14569/IJACSA.2020.0110221

[8] Luh-Maan Chang,Po-Han Chen and Heng-Kuang Shen from Department of Civil Engineering, National Taiwan University ," Recognition for color rust images based on artificial neural network," *Automation in Construction* vol. 90, pp. 178-187, June 2018.

[9] Muhamad Farid Mavi, et. all, "Mango ripeness classification system using hybrid technique*," Indonesian Journal of Electrical Engineering and Computer Science*, vol. 14, no. 2, pp. 859~868, May 2019.

[10] Sachin B. Jadhav, Vishwanath R. Udupi, Sanjay B. Patil, "Soybean leaf disease detection and severity measurement using multiclass SVM and KNN classifier," *International Journal of Electrical and Computer Engineering (IJECE)*, Vol. 9, No. 5, pp. 4077~4091, October 2019. ISSN: 2088-8708, DOI: 10.11591/ijece.v9i5.pp4077-4091

[11] Ghaiwat SN, Arora P. Cotton leaf disease detection by feature extraction. In: Research advances in the integration of big data and smart computing, Oct 13:892015

[12] S. Batmavady, S. Samundeeswari, Detection of Cotton Leaf Diseases Using Image Processing, *International Journal of Recent Technology and Engineering (IJRTE).* ISSN: 2277-3878, Volume-8 Issue-2S4, July 2019.

[13] Kuldeep Singh, Satish Kumar2, Pawan Kaur, "Automatic detection of rust disease of Lentil by machine learning system using microscopic images*," International Journal of Electrical and Computer Engineering (IJECE),* vol. 9, no. 1, pp. 660-666, February 2019. ISSN: 2088-8708, DOI: 10.11591/ijece.v9i1.pp660-666

[14] Saadat Nasehi, Hossein Pourghassem, "A Novel Effective Feature Selection Algorithm based on S-PCA and Wavelet Transform Features in EEG Signal Classification," *IEEE 3rd International Conference on Communication Software and Networks,* 2011.

[15] C. Lv et al., "Levenberg–Marquardt Backpropagation Training of Multilayer Neural Networks for State Estimation of a Safety-Critical Cyber-Physical System," *in IEEE Transactions on Industrial Informatics*, vol. 14, no. 8, pp. 3436-3446, Aug. 2018

[16] https://www.cse.unsw.edu.au/~cs9417ml/MLP2/BackPropagation.html, Accessed 06 June 2019.

[17] Moumita Chanda, Mantosh Biswas, "Plant Disease Identification And Classification Using Back-Propagation Neural Network With Particle Swarm Optimization," *Proceedings Of The Third International Conference On Trends In Electronics And Informatics (ICOEI 2019)*, ISBN: 978-1-5386-9439-8, 2019.

[18] *Beatriz* A. Garro and Roberto A. Vázquez2, "Designing Artificial Neural Networks Using Particle Swarm Optimization Algorithms," *Hindawi Publishing Corporation,Computational Intelligence and Neuroscience*, vol. 2015, Article ID 369298, 20 pages, 2015.

[19] Moumita Chanda, Mantosh Biswas, "Plant Disease Identification And Classification Using Back-Propagation Neural Network With Particle Swarm Optimization*," Proceedings Of The Third International Conference On Trends In Electronics And Informatics (ICOEI 2019)*, ISBN: 978-1-5386-9439-8

[20] *Plant* Village Dataset, https://plantvillage.psu.edu/diseases, Accessed 02 September 2019

[21] *Forestry* Dataset, https://www.forestryimages.org/index.cfm, Accessed 08 October 2019

[22] *UCI* dataset, https://archive.ics.uci.edu/ml/datasets/leaf, Accessed 12 October 2019

[23] Ramakrishnan M., "Groundnut leaf disease detection and classification by using back probagation algorithm." *In: IEEE international conference on communications and signal processing (ICCSP),* April, pp 0964-0968, 2015.

[24] *Pujari* JD, Yakkundimath R, Byadgi AS (2016)" SVM and ANN based classification of plant diseases using feature reduction technique*." International Journal of Interactive Multimedia and Artificial Intelligence*, ISSN-e 1989-1660, vol. 3, no. 7, pages. 6-14, 2016

[25] Mohan KJ, Balasubramanian M, Palanivel S., "Detectionand recognition of diseases from paddy plant *leaf* images," *International Journal of Computer Applications (0975-8887)*, vol. 144, no.12, June 2016.

BIOGRAPHIES OF AUTHORS

Kapilya Gangadharan is a Research scholar in Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Chennai. She is currently working as Utility Analyst at Fidelity National Information services, Durham, NC, USA with total of 6 years' teaching experience as Assistant Professor and 2 years' experience in IT. She received B. Tech degree in Information Technology from CSI College of Emgineering, Ketti, Nilgiris and M.E in Computer Science and Engineering from Periyar Maniammai University, Thanjavur. Her current research includes Digital Image Processing, Machine learning, Precision farming, Plant Technology and Production Engineering.

G Rosline Nesa Kumari received her Ph.D degree in Computer Science and Engineering at Dr MGR University Chennai, and received her M.E. Degree from Sathyabama University Chennai in 2005. She is having Sixteen Years of teaching experience. At present she is working as a Professor, Department of Computer Science and Engineering in Shadan women's college of Engineering and Technology, Telangana. She published sixty-five research publications in various International, National Conferences and Journal. She is a life member of Indian Science Congress Association (ISCA), IAENG, IET, CSI (Life Member), Red Cross. Her research interest includes Image processing, Steganography, Network security and Information Security includes Image processing, Digital Watermarking and Security.

D.Dhanasekaran received his PhD degree with the specialization of VLSI reconfigurable circuits. He had 29 years of teaching experience. He is working as a Principal in Saveetha school of Engineering. Saveetha Institute of Medical and Technical Sciences, Chennai. He published 50 Scopus indexed papers and two patents. His area of research includes VLSI circuits and IoT.

K. Malathi is an Associate Professor in the Department of Computer Science and Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Chennai with 11.6 Years of teaching experience and 1.6 years of industry experience. She graduated her B.Tech from Anna Univeristy, M.E from Sathyabama University, and Ph.D degree in computer science in 2018 from Saveetha Institute of Medical and Technical Sciences, Chennai. Her research interests include image processing, datamining, Medical Image Processing. She is life member in CSI, IEEE Computer Society, and IET. She has published around 40+ research papers in reputed journals and 4 International conferences. She has received most Enthuastic Faculty Award in women's day.