

PLANT DISEASES CLASSIFICATION USING FEATURE REDUCTION, BPNN AND PSO

Moumita Chanda¹ and Mantosh Biswas²

Department of Computer Engineering, NIT Kurukshetra, India^{1,2}
moumitachanda12@gmail.com¹ and mantoshb@gmail.com²

ABSTRACT

Agriculture is the culture of land and rearing of the plants to provide food for the nourishment and enhancement of life. In India, it is one of the fundamental financial sources; various sorts of plants are cultivated each year. There are various micro-organisms that cause many plant diseases and impede normal plant growth. That is the reason from long ago which led researchers to search for new methods of classification of plant diseases. Although there are different neural networks have already been used for plant disease classification, but using these methods alone do not create the best trade-off between time and precision. So to remove this constraint, we proposed a method for plant disease classification based on Back-propagation Neural Network (BPNN) and Particle Swarm Optimization (PSO). Now we have added some more data to our dataset and applied principal component analysis to reduce the number of total features and on these features we have applied BPNN with PSO. We first train neural network using back-propagation and then we further use PSO to get more optimized weights or fine-tune the parameters of neural network. In our experiment, we have used images of leaves that are infected by various bacterial and fungal diseases: *Alternaria alternata*, Anthracnose, Bacterial blight, Bacterial leaf scorch, *Cercospora* leaf spot, and Downy mildew, and our proposed method achieves approximately 96.42 % precision.

Keywords: Image, Clustering, Feature reduction, Principal component analysis, Back-propagation neural network, Particle swarm optimization

INTRODUCTION

There are different micro-organisms that infect plants with several diseases. These diseases reduce the growth and quality of the plants. There are different symptoms for every disease which identifies it. It may be shape, color or texture; according to these features, one can distinguish the diseases very easily with naked eyes. When it comes to machine the same concept is used to take the numerical values of these features to classify the diseases. Fungus, bacteria, and virus are the one which causes several diseases to the plants. At the first stage of infection, we can stop the severity of the disease if it can be detected. Otherwise, it gradually spreads all over the plant and makes it uncontrollable. Rapid advancement in technology makes it possible to protect the plants by identifying diseases in the initial stage.

Different micro-organisms are responsible for different plant diseases: bacteria, virus, fungus, and nematode. Every micro-organism has its unique nature and so has its unique way to infect and spread over the plants. Some common fungal plant diseases are *Alternaria alternata*, Anthracnose, Downy mildew, and *Cercospora* leaf spot [Plant Natural]. *Alternaria alternata* is a fungus which causes the black spot on leaves and stems of various plants. They

infect the crops at the time of cold storage and thus most of the times seen at the marketing time, making a huge loss. Anthracnose is a disease caused by fungi, genus *Colletotrichum*. These pathogens produce dark, water-soaked lesions on stem, leaf, and fruit. The lesion is covered by spores with slight pink color and masses in moist. Downy mildew infects many plants and produces white and slight yellow patches on upper surfaces of the leaves at rainy season. They vanish when sunny weather arrives. *Cercospora* leaf spot is caused in older leaves first and then it spreads to the newer one. Lesions are light gray colored with brown to the purple border and tiny black dots. Two common bacterial diseases are Bacterial blight and Bacterial leaf scorch. Bacterial blight is caused by the bacterial pathogen *Xanthomonas campestris* pv. *translucens* (syn. *X. translucens*). Lesions are initially pale green spots which eventually become water soaked and in later stages appear as dry dead spots. Bacterial leaf scorch (BLS) is a disease which affects many crops, caused mainly by the xylem-plugging bacterium *Xylella fastidiosa*. It is characterized by red-brown spot at the margin of the leaf which eventually increases and spreads over the whole leaf. We have taken leaves affected by these six diseases to detect them using their textural and physical dissimilarity.

Nowadays researchers are constantly trying to find out new techniques which beat the previous one. The world is being advanced every single day with new innovative ideas. In this technical era, every research activity makes our life easy, and effortless. Agriculture has also not been deprived of that. In different areas of agriculture technical superiority comes into play: plant disease detection, identifying the nutritional deficiency in the plants, controlling the growth of weeds by recognizing them in the first place, and recognizing the species of the plants from their leaves. Disease detection is time consuming, tiresome, and monotonous work when it is done manually. Replacing this work by machine comprises of acquisition of images, segmentation, feature selection, and classification. At first, some sample images are collected to train the machine. Then, images are segmented to identify the affected areas; features are extracted from the affected regions and classified by the trained machine.

In this paper, we have proposed an automatic disease classification method having six steps. All the steps have important role in disease detection. At first, the image is pre-processed through resizing, enhancing the contrast, masking the green pixel and transforming the color model. Then, the picture is fragmented to obtain the ailing segment of the leaf which is utilized to extract and reduce the features and to characterize the plant disease utilizing our proposed method which utilizes BPNN and PSO. Researchers have discovered multiple method of classification of plant disease based on back-propagation neural network from a long time ago. In order to classify plant diseases, the back-propagation algorithm is very efficient and effective, but sometimes it converges to a local optimal solution that debarks it from reaching the optimal global solution in optimal time. Neural Network has been trained using back-propagation algorithm and PSO has been only used for fine-tuning the parameters after the training is completed. So the initialization of the weights is done using the weights achieved by the back-propagation training. PSO is a tool that searches with reduced time and cost throughout the search space. Therefore, using both BPNN and PSO combine is a good cost-accuracy trade-off that is observed in our experimental outcome.

Our paper is distributed in different sections. "LITERATURE REVIEW" section shows the existing methods for plant disease identification and classification. "PROPOSED WORK" illustrates the details of our proposed method. Section "RESULTS AND DISCUSSION"

gives the results of our experiment. “CONCLUSION” presents the conclusion from the results and discussion which is followed by references in section “REFERENCES”.

LITERATURE REVIEW

Plant disease identification and classification is an application that researchers have been working on for a long time. Different methods with different classifiers: NNs, SVM, KNN, and deep learning based methods have been proposed by the researchers for identifying plant diseases.

Ramakrishnan. M et al. (2015) proposed a back-propagation neural network based method to classify groundnut leaf diseases: Alternaria leaf blight, Cercospora, Phaeoisariopsis personata, and Cercosporidium Personatum. Here, the RGB to HSV color model transformation and green pixel masking are used. After that, features are extracted and used for training of NN using the back-propagation algorithm. Radial basis function neural network is another ANN with three layers: one input layer, one hidden layer with different activation functions (non-linear), and one output layer. Though it has an almost similar structure with a back-propagation neural network, its training and testing characteristics are very different from BPNN. The radial basis function neural network is trained faster than the back-propagation neural network with a very high time difference. Toran Verma et al. (2017) used discrete wavelet features and radial basis function network for identifying five paddy diseases: Brown spot, Leaf blast, Stem borer, Sheath blight, and Panicle blast. In this work, DWT has been used to extract the features and RBFNN has been used to classify and test its performance. Shiddhart Singh Chouhan et al. (2018) proposed a solution for identifying and classifying different fungal diseases with the help of region growing method used for feature extraction and RBFNN with BFO for classifying the diseases. Here, average specificity and sensitivity of the proposed method are more compared to GA and KM methods. Probabilistic neural network (PNN) is the network which has four layers. It has been used for lots of application and performs well. It is faster than the backpropagation neural network and is affected very less by unimportant outliers. The problem with this network is its complicated structure and consumption of memory. Pushpa Rani M.K. (2015) proposed a solution which uses green pixel masking, extraction of geometric features, and PNN to identify five types of plant diseases. Pooja Kulinavar et al. (2017) used KMC and multiclass SVM to classify four different plant diseases: Alternaria alternate, Cercospora Leaf Spot, Anthracnose, and Bacterial Blight. Vijay Singh et al. (2017) proposed an automatic plant disease recognition system where GA is used for segmentation. Uses of the proposed algorithm for segmentation increases the performance in each case of using Minimum Distance Criterion with KMC and SVM for classification. Megha S. et al. (2017) used fuzzy c-means clustering for segmenting the image and SVM classifier for classification. Here, Color correlogram for extracting color features, and SGDM and Otsu method have been used for texture and shape features respectively. Fuzzy c-means has been proved to give a better result when used for segmentation in place of k-means clustering in many cases. Pranjali B. Padol et al. (2016) proposed an automatic disease detection system using Gaussian filtering-means segmentation and SVM classification. Jagadeesh D. Pujari et al. (2016) proposed a solution where huge numbers of color and texture features are first selected and then they are reduced in number. Then taking these reduced features SVM classifier and ANN are separately used for classification. The results have been compared where SVM classifier gives more accurate result than ANN. Iqbaldeep Kaur et al. (2016) proposed a disease detection solution which uses histogram equalization for contrast enhancement, KMC and SVM classifier. Yogesh Dandawate et al. (2015) proposed a futuristic decision support system for farmers which uses

SIFT to recognize plant species using shape. The SVM classifier is used to differentiate disease-free and diseased leaves (soybean) with an accuracy of 93.79%. KNN classifier is a simple algorithm that basically checks the Euclidean distances between the unknown data point and the input data points. The k number of input data points which are closer to the unknown data point are considered to choose its class. The classes of the k number of data points are counted and the maximum appeared class is considered as the class of the unknown point. The performance of KNN classifier is highly dependent on the value of k. Thus, there are different algorithms which approximate the value of k very efficiently. The simplest one is k-fold cross-validation. Gautam Kaushal et al. (2017) used GLCM features and K-nearest neighbor classifier to identify plant diseases and the features used are contrast, energy, entropy, homogeneity, and others. Viraj A. Gulhane et al. (2014) gave a solution to classify cotton diseases: Blight, Gray Mildew, Leaf Necrosis, Magnesium Deficiency, and Alternaria using PCA and KNN. Haiguang Wang et al. (2012) used seven sets of features to identify grape diseases: Downy mildew, and Powdery mildew and wheat diseases: wheat stripe rust, and wheat leaf rust using the back-propagation neural network. Here, images are converted to LAB color format and segmented using k-means clustering. Twentyone color features: mean of gray values of R, G and B components, mean of gray values of H, S and V component etc., four shape features: area, complexity, perimeter and circularity, and twentyfive texture features: contrast, entropy, correlation, etc. are extracted, which are divided into seven different combinations. These seven combinations of inputs are processed through principal component analysis and given to BPNN individually to identify the plant diseases. Reza Ghaffari et al. (2010) used three classifiers: Radial Basis Function, Learning Vector Quantization, and Multilayer Perceptron to classify tomato diseases. Here, data are collected by electronic notes, reduced by PCA, and then segmented by KMC and FCC. Here, healthy and diseased tomato leaves are classified by three classifiers: multilayer perceptron, learning vector quantization, and radial basis function. Camilo Pulido et al. (2017) showed uses of PCA and SVM to identify weeds, and vegetables from outdoor crop images. Different texture features are extracted and used in this work which includes: autocorrelation, contrast, correlation, energy, dissimilarity, entropy, homogeneity, and variance. These features are then reduced using PCA and used to classify using SVM. Performance criteria used in this work are true positive rate, false negative rate, true predicted value, and negative predicted value. Stephen Gang Wu et al. (2017) proposed a solution to recognize thirty types of leaf images using PNN. In this work, first, the RGB image is converted to the grayscale image which is then further converted to a binary image. After that rectangular averaging filter and laplacian filter are used to remove noise and to get the boundary of the leaf respectively. To make the boundary as a black outline all the pixels are swapped between zero and one. Then, twelve morphological features are extracted: aspect ratio, form factor, narrow factor, smooth factor, rectangularity, etc. The Principal component analysis is used to get independent orthogonal features which are then used to identify the type of the leaf. Viraj A. Gulhane et al. (2014) gave a solution to classify cotton diseases: Blight, Gray Mildew, Leaf Necrosis, Magnesium Deficiency, and Alternaria using PCA and KNN. Here, the color change of the green channel is taken as the main features. Cosine distance used to measure distances between training and testing vectors and based on the distances the leaves are classified to the right diseases.

We have seen from the literature survey above that NN is very effective in the identification of plant diseases. NN has a fixed size which represents an advantage over SVM. This consists of a different number of different size hidden layers depending on the number of features, dataset, connection edge's weights, and biases. On the other hand, there are some problems with NN: Local minima and overfitting. To overcome these problems, we have used PSO

after training of NN with back-propagation for classifying diseased leaf images in our proposed method. We have described our proposed method in detail in the next section.

PROPOSED WORK

There are six steps in our proposed method which are used sequentially to identify plant diseases. Images are obtained from a dependable source in the first step. There are different websites having the collection of high-quality diseased plant images. Images can be gathered either from these sources, or they can be clicked by the camera to obtain new sets of images. In the second step collected images are pre-processed using resizing, green pixel masking, contrast adjustment, and conversion of the color model. Pre-processing makes the image more visible, clean and suitable for further processing. In third step image is segmented into different dissimilar regions with respect to change of color, using k-means clustering. In the fourth step, important features are extracted from the segment consisting of diseased pixels. In fifth step number of features are reduced using PCA. In the final step, image is classified into different diseases using NN. Here we have used back-propagation algorithm for training the NN and PSO for further optimization of NN weights. PSO makes the plant disease detection faster than the other existing methods by increasing the convergence rate. Following are the detail description of each and every step of our proposed method.

3.1. Step 1: Image Acquisition

Image acquisition is an important step in the identification of disease, since output varies depending on image quality. In our proposed solution we took images of five types of diseased leaves from an online source [www.forestryimages.org] that were already used to introduce successful application for disease detection. The dataset consists of five types of leaves: *Alternaria alternata*, Anthracnose, Bacterial blight, Bacterial leaf scorch, *Cercospora* leaf spot, Downy mildew and Healthy Leaves (Figure 3).

3.2. Step 2: Image Pre-processing

Image pre-processing is an important step in the identification of disease to eliminate unnecessary details in the image of the leaf. Image resizing helps to reduce the processing time in the later disease detection stages. In this process the image is transformed to the dimensions 256 * 256 and the image contrast is increased. Masking green pixels helps eliminate the leaves' healthy areas. Thereby the pixels whose green component value is greater than the other two components: red and blue are eliminated. In Figure 1(b) the image of the leaf has been shown after removal of the green pixels. The picture with RGB format is transformed in color space transformation to CIE L*a*b color model which is device-independent. CIE L*a*b color space only carries color information in a & b portion which helps reduce the image dimension and make the image suitable for segmentation.

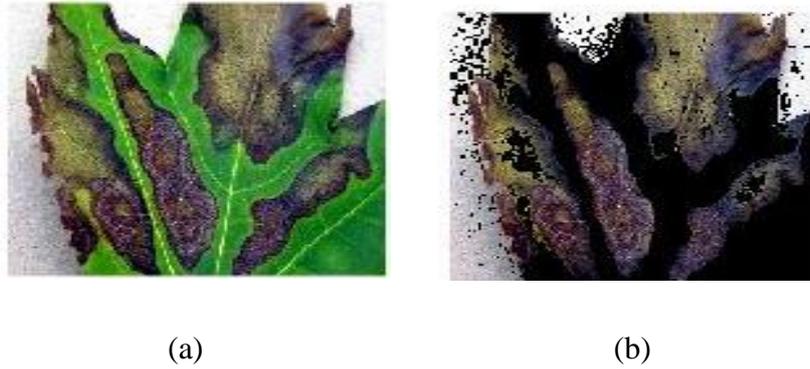


Figure. 1. Leaf image (a) Anthracnose and (b) After high green pixel elimination

3.3. Step 3: Image Segmentation

Segmentation is the process by which the image is separated into different regions having the same characteristics. It helps to get more in number the area that consists of the diseased pixels. One of the best segmentation techniques is k-means clustering which divides the image into k number of clusters based on its color. In our proposed method we use KMC to divide the image into three clusters which has been shown in Figure 2. Leaf image is segmented based on the various colors present in the image and one of these clusters is selected as the area of interest. Here cluster with more pixels affected by the disease is taken and forwarded for further processing.

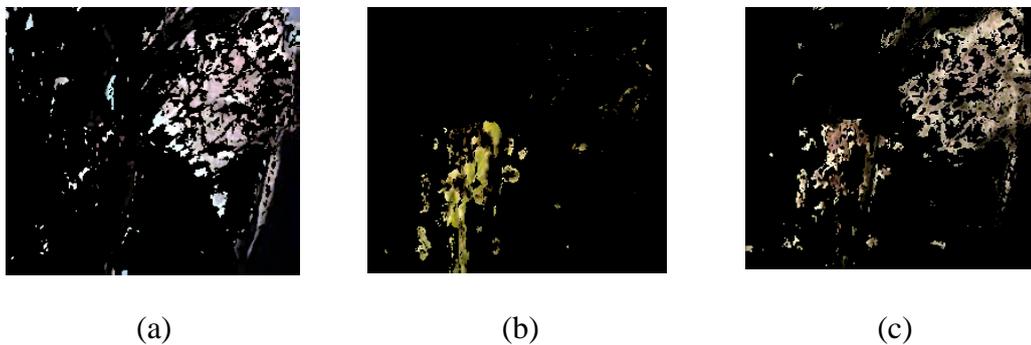


Figure. 2. Segmented image of Cercospora leaf spot (a) Cluster 1 (b) Cluster 2 and (c) Cluster 3

3.4. Step 4: Feature Extraction

For plant disease detection, feature extraction is used to remove the infected portion of the leaf image. The leaf disease has certain unique characteristics which help to identify it. Features used for better identification of plant disease are: color, shape, and texture. Such features are the backbone of the disease detection process, and efficiency depends on the proper selection of these features. In our proposed method we use different texture features: Contrast, Correlation, Entropy, Energy, and Homogeneity which are extracted from gray level co-occurrence matrix obtained from GLCM function. Contrast represents the number of locally present variations in an image. The amount of uniformity present in the image is determined by energy. Local homogeneity tests the correlation between the distribution of GLCM elements and their diagonal components. The sum of dissimilarity present in the GLCM matrix is determined by entropy. Correlation tests the joint probability of two pixel

values appearing in an image. Mean, Root Mean Square (RMS), Variance, Smoothness, Kurtosis, Standard Deviation, Skewness and Inverse Differential Momentum (IDM) are other statistical features that we used in our method.

3.5. Step 5: Feature reduction

Here we use principal component analysis which helps to choose the components which are uncorrelated, independent and has a huge contribution in disease detection. PCA does this using the orthogonal transformation of the original correlated features to a reduced set of uncorrelated features. The components are chosen in the decreasing order of variability. The components which have negligible variance after discriminant analysis are eliminated. Thus, principal component analysis reduces the number of features without losing any information and increases the performance of the application. After extracting all the above features we reduce the number of features which takes twelve out of thirteen features. In the next step, we classify the leaves using back-propagation neural network and particle swarm optimization.

3.6. Step 6: Classification

Classification helps to classify each type of leaves based on their texture, color and shape features. In our proposed method classification is done using back-propagation and PSO. At first, back-propagation is used to train the feed forward neural network and then NN connection weights are further optimized using particle swarm optimization.

3.5.1. Feed-forward Neural Networks

Neural networks are good in approximating any non-linear function. There are different algorithms to train the neural network. The neural network has long been proven to be a good classifier and has been applied successfully in many complex and major classification problems [Bacha Rehmam et al. (2015)] [M. Hariharan et al. (2010)]. A feed forward neural network with one input layer with thirteen nodes, two hidden layers with thirteen nodes each and one output layer with three nodes is used here in our proposed method. Every input node represents one feature and no. of output nodes depends on the no of classes. Neural network is trained using back-propagation algorithm which consists of two phases: feed-forward phase and backward error correction phase.

(i) Feed-Forward Phase

In the feed forward phase inputs are propagated across hidden layers and the output is calculated by multiplying the weight vector (between the hidden and output layer) with the input vector (hidden layer). After inputs are propagated in the forward direction through all NN layers, the difference between the target output and the desired output or error is calculated and this error goes backward in the backward phase.

(ii) Backward phase with Conventional Back-propagation

Back-propagation algorithm has a major constraint that hidden layer neuron activation function should be differentiable. In our proposed method learning is achieved with the aim of minimizing the mean square error (MSE). Here, the target output and hidden layer error are first calculated and then weights are updated based on these errors.

There are some parameters whose values should be chosen properly to have good performance or fast learning. The most important one is learning rate. A too low or too high

value of this parameter decreases the accuracy of plant disease detection. Its value should be chosen optimally to have fast global convergence. When there are many local and global optima then variable learning rate is good. In our experiment learning rate has been taken as 0.15 to have a good trade-off between speed and accuracy. The second important parameter is activation function which is used to get the final output from the nodes. In our experiment, we have used the binary sigmoidal function. We are using gradient descent for updating the weights when the neural network is trained using back-propagation neural network and NN weights are used to set the positions of the particles in PSO which we have described below.

3.5.2. Particle Swarm Optimization (PSO)

Using PSO, the weights are further adjusted after training is completed by BP with a predefined termination condition that is if no. of desired iterations are reached or not. PSO is an evolutionary technique which employs a number of populations to search for the optimum solution.

Initially, an arbitrary position and velocity of the particles are chosen. Then, the velocity and position are updated incrementally. There are three components in velocity equation: previous velocity, the difference between 'Pbest' & current position and the difference between 'gbest' & current position. The first component lets particles move in the direction of their preceding solution. The second component considers the particle's best solution or location until the moment of the velocity update. The third component directs the particle to the best global solution found till the time of velocity update. Every component is given certain weight according to the iteration number. So PSO's basic concept is to move particles towards best social and global solutions. At first, the population of particles is initialized taking random values for positions and velocities of particles using previously obtained weights from the neural network to set the limit. Here the weights are taken as ± 0.1 to the output weights of BPNN as the limit. 0.1 is taken because it is giving the best result for our classification problem. For each particle, the fitness function is measured and compared with the global best cost. Here we use negative accuracy as the fitness function to make it a minimization problem. If the fitness function value is less than the global best, then global best position and cost are updated. Each particle's velocity, V_i and position, X_i are updated by the following equations, (1) and (2), respectively.

$$V_i(t+1) = W * V_i(t) + c1 * rand1 * (P_{ibest} - X_i(t)) + c2 * rand2 * (G_{ibest} - X_i(t)) \quad (1)$$

$$X_i(t+1) = X_i(t) + V_i(t+1) \quad (2)$$

Here, $V_i(t+1)$ is the current velocity of the particle and $V_i(t)$ is the velocity of the previous time step. $c1$ and $c2$ are positive constants used to add different weighting to the social and global components, respectively. We set the values of $c1$ and $c2$ by equations (3) and (4), respectively. $rand1$ and $rand2$ are random values between 0 and 1. $X_i(t+1)$ is the current position whereas, $X_i(t)$ is the previous position of the particle.

$$c1 = (c11 - c12) * (it / nPop * 10) + c12 \quad (3)$$

$$c2 = (c22 - c21) * (it / nPop * 10) + c21 \quad (4)$$

Here, $nPop$ is the number of particles, and $c11$ and $c21$ are the lower limits and $c12$ and $c22$ are higher limits of social and global coefficients. In our experiment we have taken $nPop =$

100, $c_{11} = c_{21} = 0.5$ and $c_{12} = c_{22} = 2.5$. The fitness function is calculated and compared with 'Pbest'. If current position's cost is less than the previous best position, then 'Pbest' is updated. Then current 'Pbest' of each particle is compared with 'gbest' and if it is less than 'gbest' then, the global position is updated. Step (iii) is continued until a predefined criterion that is the maximum number of iterations is met. When we use PSO one main concern is to set the parameter's values appropriately. There are following parameters: inertia weight, the coefficient of social component, coefficient of the global component, and the velocity (i.e., max and min) in PSO. Changing one of its value changes the performance of the PSO drastically. Here, the performance of the PSO basically refers to the convergence rate of PSO. After a different hit and trial attempt, the values of the parameters have been set to the following numerical values to achieve the best performance. We have used -4 and 4 for Vmax and Vmin, respectively and 0.9 for inertia weight with a damping ratio of 0.99.

RESULTS AND DISCUSSION

Experimentally, we have analysed the contribution of six different methodologies after applying them on our considered dataset: Haiguang Wang et al., Reza Ghaffari et al., Camilo Pulido et al., Stephen Gang Wu et al., and Viraj A. Gulhane et al. to detect six types of diseases namely *Alternaria alternata*, Anthracnose, Bacterial blight, Bacterial leaf scorch, *Cercospora* leaf spot, Downy mildew, and healthy leaves shown in Figure 3. For the experiment we have divided the total set of the images into two parts: one set has been used for training which is 60% of the total dataset and another set has been used for testing that is 40 % of the dataset. For each type of leaf, sixty images have been used for training and forty images for testing. Total two hundred eighty images have been used for testing and comparing the performance of our proposed with other five above methods. We have implemented the considered methods using MATLAB R2016a running at Core I3, 2.20 GHz processor with 4GB. The criteria used to measure the correctness of the methods is accuracy.

Table 1 represents the result of our proposed method for each and every category of diseases of the considered dataset. Here, we observe that for different classes different accuracy rate has been achieved.

In Table 2, we see the performance comparison for plant disease detection between our method and other classification based existing methods: Haiguang Wang et al., Reza Ghaffari et al., and Camilo Pulido et al., Stephen Gang Wu et al., Viraj A. Gulhane et al.'s method based on classification accuracy. Here, we have considered average accuracy of each method over 5 iterations and accuracy is calculated as the percentage of correctly classified images over total a number of taken testing images. From the above study and result BPNN has been proved as the best NN based classifier that can identify the considered six plant diseases accurately but BPNN takes the highest time among the five classifiers: BPNN, SVM, KNN, PNN, and RBFNN. To keep the advantage of BPNN and reduce the time complexity of the application we have used particle swarm optimization with BPNN in our proposed method. Our proposed method's accuracy is 96.42 whereas Haiguang Wang et al.'s BPNN based method's accuracy is 95.71. Using our considered dataset Reza Ghaffari et al.'s RBFNN based method's accuracy is 93.93%, and Camilo Pulido et al.'s SVM based method's accuracy is 82.5%, Stephen Gang Wu et al.'s PNN based method's accuracy is 75.36%, Viraj A. Gulhane et al.'s KNN based method's accuracy is 64.29%. Comparative analysis in table 2 shows that our method gives the best result among the six methods using our considered dataset.

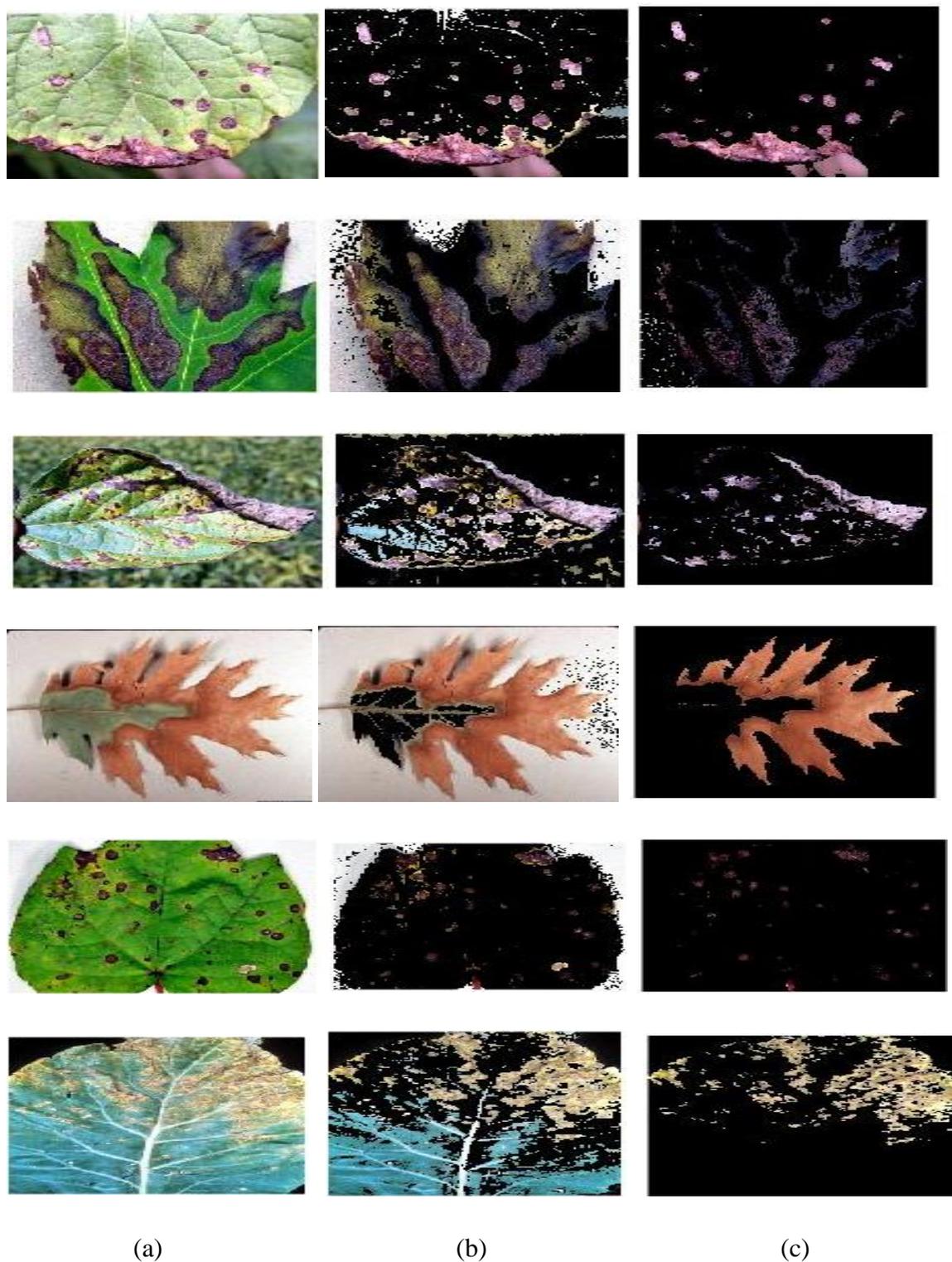


Figure 3. Six types of diseased leaf images: *Alternaria alternata*, Anthracnose, Bacterial blight, Bacterial leaf scorch, *Cercospora leaf spot*, and Downy mildew (sequentially): (a) original (b) after high green pixel elimination and (c) segmented, respectively.

Table 1. Results of the proposed method for considered leaf images

Type of leaf image (Diseased/Healthy)	Number of images for testing	Number of correctly classified images	Accuracy (%)
Alternaria Alternata	40	39	97.5
Anthracnose	40	38	95
Bacterial Blight	40	39	97.5
Bacterial Leaf Scorch	40	40	100
Cercospora Leaf Spot	40	36	90
Downy Mildew	40	38	95
Healthy	40	40	100

Table 2. Comparative results of accuracy for identification of six considered leaf diseases using six considered method

Methods	Number of images for testing	Accuracy
Proposed method	280	96.42
Haiguang Wang et al.	280	95.71
Reza Ghaffari et al.	280	93.93
Camilo Pulido et al.	280	82.5
Stephen Gang Wu et al.	280	75.36
Viraj A. Gulhane et al.	280	64.29

CONCLUSION

Principal component analysis chooses and selects the uncorrelated and independent features which are used for the classification and particle swarm optimization is a fast evolutionary algorithm in computational terms that has both theoretical and practical significance. Using PSO with BPNN after reducing the number of features using PCA gives a stable and accurate classification. In this paper, we have increased our dataset and applied our proposed classification method BPNN with PSO on six plant diseases. We have also seen the comparison of our method with other six considered methods on our dataset. In future, we

want to continue our research for better accuracy and classification despite of having increased number of samples in the dataset.

REFERENCES

- Bacha Rehman, Zahid Hali, Ghulam Abbas, Tufail Muhammad, "Artificial neural network-based speech recognition using DWT analysis applied on isolated words from oriental languages", *Malaysian Journal of Computer Science*, vol. 28, issue 3, pp. 242-262, 2015
- B.M. Cooke, D.G. Jones and B. Kaye, *The Epidemiology of Plant Diseases*, New York: Springer, 2006
- Camilo Pulido, Leonardo Solaque, and Nelson Velasco, "Weed recognition by SVM texture feature classification in outdoor vegetable crop images", *Engineering and Research*, vol. 37, no. 1, pp. 68-74, 2017
- D. J. Pujari, R. Yakkundimath, and A. S. Byadgi, "SVM and ANN based classification of plant diseases using feature reduction technique" *International Journal of Interactive Multimedia and Artificial Intelligence*, vol. 3, no. 7, pp. 6-14, 2016
- Gautam Kaushal, Rajni Bala, "GLCM and KNN based Algorithm for Plant Disease Detection", *International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering*, vol. 6, issue 7, pp. 5845-5852, 2017
- Haiguang Wang, Guanlin Li, Zhanhong Ma, and Xiaolong Li, "Image Recognition of Plant Diseases Based on Back-propagation Networks", *International Congress on Image and Signal Processing*, pp. 894-900, 2012
- John William Orillo, Jennifer Dela Cruz, Leobelle Agapito, Paul Jensen Satimbre and Ira Valenzuela, "Identification of Diseases in Rice Plant (*Oryza Sativa*) using Back Propagation Artificial Neural Network", *IEEE International Conference on Humanoid, Nanotechnology, Information Technology Communication and Control, Environment and Management, Palawan, Philippines*, pp. 1-6, 2014
- Kaur, G. Aggarwal, and A. Verma, "Detection and classification of disease affected region of plant leaves using image processing technique", *Indian Journal of Science and Technology*, vol. 9, no. 48, pp. 1-13, 2016
- Lucas, G.B., C.L. Campbell and L.T. Lucas, *Introduction to Plant Diseases*, Westport, CT: AVI Publishing, 1985
- Maicon A. Sartin, Alexandre C.R. Da Silva and Claudinei Kappes, "Image segmentation with artificial neural network for nutrient deficiency in cotton crop", *Journal of Computer Science*, vol. 10, issue 6, pp. 1084-1093, 2014
- Moumita Chanda and Mantosh Biswas, "Plant disease identification and classification using Back-Propagation Neural Network with Particle Swarm Optimization", *3rd International Conference on Trends in Electronics and Informatics*, pp. 1-8, 2019
- M. Hariharan, M. P. Paulraj, Sazali Yaacob, "Time-domain features and probabilistic neural network for the detection of vocal fold pathology", *Malaysian Journal of Computer Science*, vol. 23, issue 1, pp. 60-67, 2010
- P. B. Padol, and A. A. Yadav, "SVM classifier based grape leaf disease detection", *Conference on Advances in Signal Processing*, pp. 175-179, 2016
- Pooja Kulinavar, Vidya I. Hadimani, "Classification of leaf disease based on Multiclass SVM classifier", *International Journal of Advance Research, Ideas and Innovations in Technology*, vol. 3, issue 4, pp. 321-327, 2017

- Pushpa Rani M.K., "Detection and Analysis of Plant Leaf Diseases using Image Processing", *National Conference Electronics, Signals, Communication and Optimization*, pp. 8-13, 2015
- Ramakrishnan M., Sahaya Anselin Nisha A., "Groundnut leaf disease detection and classification by using back propagation algorithm", *International Conference on Communication and Signal Processing*, 2015, pp. 964-968
- Reza Ghaffari, Fu Zhang, Daciana Iliescu, Evor Hines, Mark Leeson, Richard Napier, and John Clarkson, "Early Detection of Diseases in Tomato Crops: An Electronic Nose and Intelligent Systems Approach", *International Joint Conference on Neural Networks*, pp. 1-6, 2010
- SaharMansourian, Ebrahim IzadiDarbandi, Mohammad HassanRashed Mohassel, MehdiRastgoo, HomayounKanouni, "Comparison of artificial neural networks and logistic regression as potential methods for predicting weed populations on dryland chickpea and winter wheat fields of Kurdistan province, Iran", *Crop Protection*, Vol. 93, 2017, pp. 43-51, DOI: 10.1016/j.cropro.2016.11.015
- Siddhart Singh Chouhan, Ajay Kaul, Uday Pratap Singh, and Sanjeev Jain, "Bacterial Foraging Optimization Based Radial Basis Function Neural Network (BRBFNN) for Identification and Classification of Plant Leaf Diseases: An Automatic Approach Towards Plant Pathology", *Institute of Electrical and Electronics Engineers Access*, vol. 6, pp. 8852-8863, 2018
- 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", *International Symposium on Signal Processing and Information Technology*, pp. 15-18, 2007
- S. Megha, C. R. Niveditha, N. SowmyaShree, and K. Vidhya, "Image processing system for plant disease identification by using FCM clustering technique.", *International Journal of Advanced Research*, vol. 3, no. 2, pp. 445-449, 2017
- S. Zhang, X. Wu, Z. You and L. Zhang, "Leaf image based cucumber disease recognition using sparse representation classification", *Computers and Electronics in Agriculture*, vol. 134, pp. 135-141, 2017
- Toran Verma, and Sipi Dubey, "Paddy Disease Recognition using Image Processing and Radial basis Function Network", *Indian Journal of Science and Technology*, Vol. 10, no. 46, pp. 1-17, 2017
- The University of Georgia - Warnell School of Forestry and Natural Resources and College of Agricultural and Environmental Sciences, Center for Invasive Species and Ecosystem Health, US Forest Service, International Society of Arboriculture, USDA Identification Technology Program, <https://www.forestryimages.org/Collections> (Last updated October 2018)
- Viraj A. Gulhane, and Maheshkumar H. Kolekar, "Diagnosis of Diseases on Cotton Leaves Using Principal Component Analysis Classifier", *India Conference*, pp. 1-5, 2014
- V. Singh, and A. K. Misra, "Detection of plant leaf diseases using image segmentation and soft computing techniques", *Information Processing in Agriculture*, vol. 4, pp. 41-49, 2017
- Y. Lu, S. Yi, N. Zeng, Y. Liu, and Y. Zhang, "Identification of rice diseases using deep convolutional neural networks", *Neurocomputing*, vol. 267, pp. 378-384, 2017
- Y.H. Dandawate, and R. Kokare, "An automated approach for classification of plant diseases towards development of futuristic Decision Support System in Indian perspective",

International Conference on Advances in Computing, Communications and Informatics,
vol. 978, no. 1, pp. 794-799, 2015

Y. Lu, S. Yi, N. Zeng, Y. Liu, and Y. Zhang, "Identification of rice diseases using deep convolutional neural networks", *Neurocomputing*, Vol. 267, pp. 378-384, 2017