ANALYSIS OF PLANT DISEASES USING EXPECTATION MAXIMIZATION DETECTION WITH BP-ANN CLASSIFICATION

BATTALA LAKSHMI THUSHARA[1], T.MAHABOOB RASOOL[2]

1 M.TECH, DECS, Bharath College of Engineering and Technology for Women, Kadapa

2 Assistant Professor, Dept OF ECE, Bharath College of Engineering and Technology for Women, Kadapa

ABSTRACT: The plant disease diagnosis is restricted by person's visual capabilities as it is microscopic in nature. Due to optical nature of plant monitoring task, computer visualization methods are adopted in plant disease recognition. The aim is to detect the symptoms of the disease occurred in leaves in an accurate way. Once the captured image is pre-processed, the various properties of the plant leaf such as intensity, colour and size are extracted and sent to classifier with Artificial Neural Network for classification of disease which the plant gets.

Keywords: ANN,EM,GLCM

1. INTRODUCTION

The Production of good quality food produce and improvement in crop yield are challenging for the papers as well as agriculturist to meet the growing demands globally. Thus, it is crucial to maximize agriculture resources and its utilizations in a sustainable manner. Therefore, for the sustainable agriculture system, use of emerging technology becomes important for significant and efficient contributions. With the implementation of these techniques, possibility to reducing errors and costs for achieving ecologically and economically sustainable agriculture is the thought of the present era [1-3]. Earlier used techniques were inefficient and time consuming for analyzing the problems and implementation of remedial measures. Diseased plants exhibit a variety of symptoms like, stunting, yellowing, wilting, twisting, reddening, browning, blighting, and other abnormalities [2-5]. Thus, accurate diagnosis is essential to diagnose and control the plant disease effectively. Until a disease is adequately diagnosed, a grower may waste time and energy as well as money to solve a problem with an unknown cause. Once a disease is diagnosed, appropriate management practices can be selected [6-8]. To overcome this problem a fast and accurate process is required, that can automatically detect the disease on the leaf. Technique such as visual detection requires significant time for visual inspection for a large cultivated area. Thus, image processing technique is proven to be an effective method as compared to visual analysis.

2. LITERATURE SURVEY

In the modern era, enhancement in the use of internet attracts the science and engineering techniques to get easy and quick solution, as, it is most efficient and effective way of communication. Therefore, Korkut UB et al. (2018) developed a web based tool for identifying pomegranate leaf disease. In the first step, feature extraction (based on colour and morphological features) was done. Thereafter, to segment the diseased part from healthy region k-mean algorithm and SVM was used. Accuracy achievement in this method was 65% [1]. Segmentation was done using genetic algorithm to distinguish diseased part and the healthy parts of the leaves. This algorithm was tested on ten species of plant i.e. Jackfruit, Banana, Mango, Sapota, Potato, Beans, Tomato, Lemon, etc. to check the accuracy of proposed algorithm. The author accomplished that the algorithm provides optimum results with less computational efforts in recognition and classification of leaf disease. Shima Ramesh A et al. (2018) proposed a machine learning technique for automatic detection and classification of Sugarcane leaf disease using image processing technique. Infected leaves were captured by using digital camera. After then, the preprocessing and segmentation was done using image histogram equalization, filtering, color transformation to detect infected parts of the leaves. Finally, SVM classifier was used for classification purpose [2].

3. PROPOSED METHOD

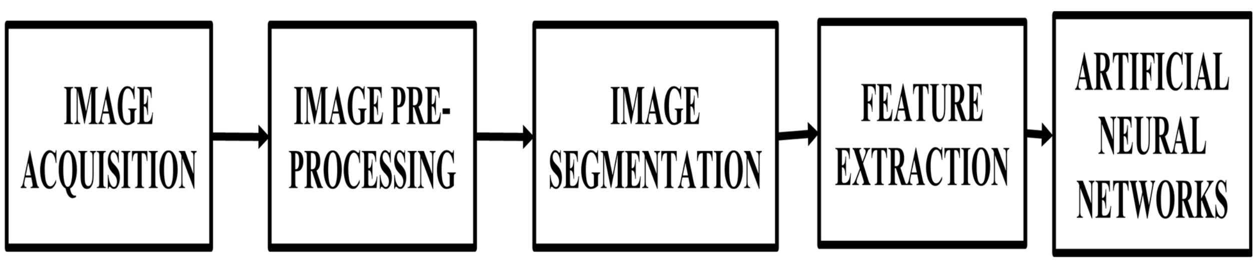

Figure 1: Proposed Block diagram

The basic operational detail of the plant disease detection and classification is represented in figure 1. The operation is explained below

IMAGE ACQUISITION: The impression of images is the primary and essential step to observe the state of the Groundnut leaf. The image imprisonment has been done through various tools and devices, such as, cameras, mobile phones and satellites. The proper estimation of RGB color pixels in an image is essential step towards successful completion of image capturing. The technical parameters of these simple, handheld devices such as light sensitivity of the photo sensors, spatial resolution and digital focusing have improved dramatically year after year. Today, nearly every person, farmer or plant pathologists carry these modern and sophisticated devices such as digital cameras together with a mobile phone or tablet computer.

IMAGE PRE-PROCESSING: The pre-processing follows the image acquisition. The acquisition of images and creating images database, pre-processing has been done. The pre-processing of created database is a preliminary step to eliminate the undesired distortion of the image and provides enhancement in features. While considering leaf of a plant, various colors have been observed. To distinguish the color of the diseased lesion from the original color of leaf, the RGB color pixels should be converted into some other pre-processing for the better perception. The reason for unacceptability of RGB is the system dependency of such pre-processing. Therefore, the improvement in the precision of color for detection of disease, the independency of pre-processing is essentially required.

IMAGE SEGMENTATION: Segmentation of an image is the process of partitioning the object (diseased spot) from its background (leaf). Different segmentation techniques are available like clustering methods, thresholding, edge detection, ANN based methods, partial differential equation based segmentation, etc. In the present paper Expectation maximization (EM) clustering technique for segmentation has been given the priority among all of the above stated techniques as shown in figure 2. The inherent advantages of k-mean clustering method are that, it works well with large data sets. The accuracy of system depends on the data sets. Therefore, this (k-mean clustering) proves to be fast, robust, easier to understand and simplest to implement. Furthermore, it may work more efficiently; if clusters are spherical (diseased spots are spherical in shape) and more in number. Increased value of the k (cluster) reduces the amount of error in the result. Current work has a value of five for k. The formations of clusters have been done based on the selection of five random points selected from the data sets. These five random points treated as centroids of each cluster. These random points attract the same intensity points (based on Euclidian distance method). This movement of the centroid happened till the same intensity cluster formed and can’t move further. The ultimate end results come in the form of diseased and healthy parts of the leaves. After segmentation, one of the diseased clusters (obtained from one or more than one cluster) has been extracted and considered for calculation of the disease area of the leaf.

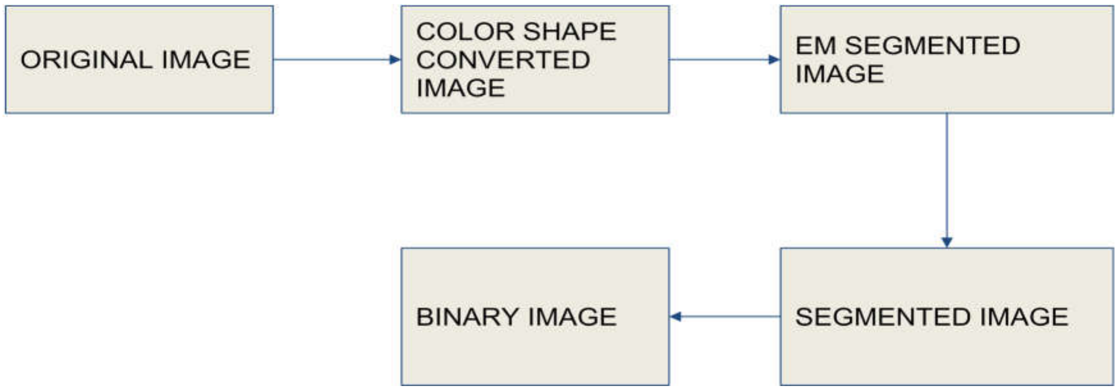

Figure 2: Proposed EM segmentation

FEATURE EXTRACTION: At this stage of the paper we calculate the Gray Level Co occurrence Matrix of an image in order to extract the set of features required for further calculations.In a statistical texture analysis, texture features were computed on the basis of statistical distribution of pixel intensity at a given position relative to others in a matrix of pixel representing image. Depending on the number of pixels or dots in each combination, we have the first-order statistics, second-order statistics or higher-order statistics. Feature extraction based on grey-level co-occurrence matrix (GLCM) is the second-order statistics that can be used to analyze the image as a texture. GLCM (also called gray tone spatial dependency matrix) is a tabulation of the frequencies or how often a combination of pixel brightness values in an image occurs. Transforming the input data into the set of features is called *feature extraction*. If the features extracted are carefully chosen it is expected that the features set will extract the relevant information from the input data in order to perform the desired task using this reduced representation instead of the full size input.

Artificial Neural Network (ANN): The key element of this paradigm is the novel structure of the information processing system. It is composed of a large number of highly interconnected processing elements (neurones) working in unison to solve specific problems. ANNs, like people, learn by example. 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. This is true of ANNs as well. The ANN consist of major three layers for training and testing operation so effective classification result will achieve. The constructions of those three layers are presented in Fig 3 and operation is explained.

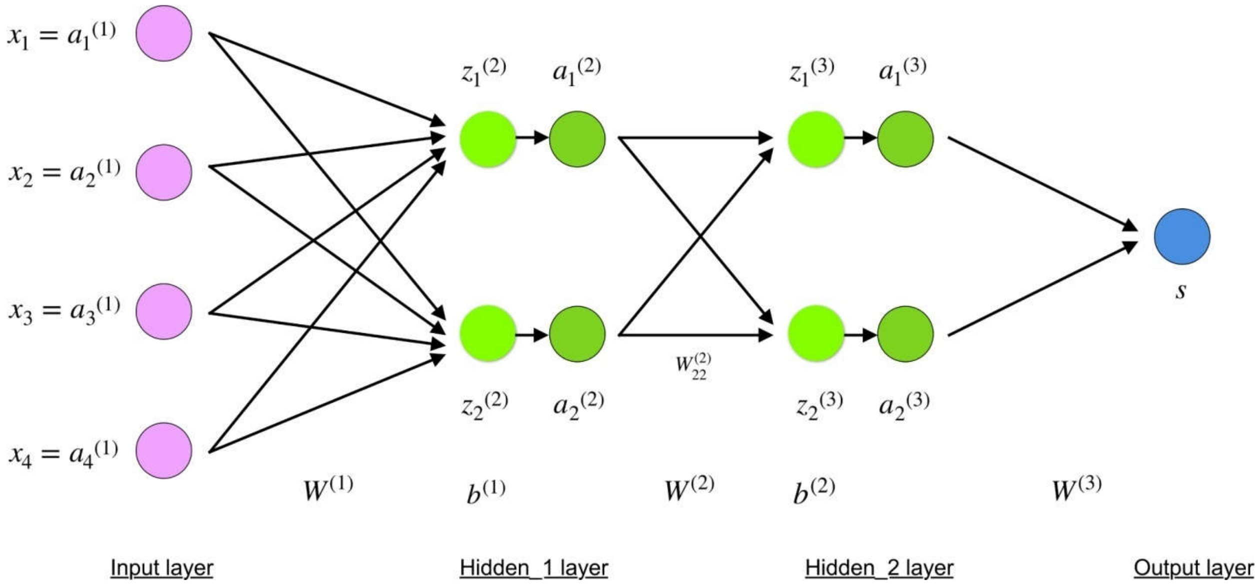

Figure 3: back propagation model of ANN

Input layer: Each neuron in the input layer represents a predictor variable. In categorical variables, *N-1* neurons are used when there are *N* number of categories. It standardizes the range of the values by subtracting the median and dividing by the interquartile range. Then the input neurons feed the values to each of the neurons in the hidden layer.

Hidden layer: This layer contains one neuron for each case in the training data set. It stores the values of the predictor variables for the case along with the target value. a hidden neuron computes the Euclidean distance of the test case from the neuron's centre point and then applies the RBF kernel function using the sigma values.

Output layer: For ANN networks there is one pattern neuron for each category of the target variable. The actual target category of each training case is stored with each hidden neuron; the weighted value coming out of a hidden neuron is fed only to the pattern neuron that corresponds to the hidden neuron's category. The pattern neurons add the values for the class they represent. The output layer compares the weighted votes for each target category accumulated in the pattern layer and uses the largest vote to predict the target category.

By the parallel process of this layers will produce the effective classification result. The Fig 3 represents the detailed architecture of Training and Testing process using ANN. In the training process, on the plant image dataset consisting of more than 1000 images are performed. In this training process, initially pre-processing operation will perform for noise removal then enhancements, segmentation operation to detect the location of disease using k means clustering then detected location features are extracted using GLCM filter. Now the detected features are applied to weight adjustment process to create a sigmoid based model function for ANN back propagation model. The weights are altered based on the chronological relationship base on the multiple plant diseases. Then using this weights an artificial intelligence based back propagation model will formed. Whenever test image applied, its GLCM features are compared with the back propagation model and classification operation will be formed. And classification results generate the type of disease with high accuracy.

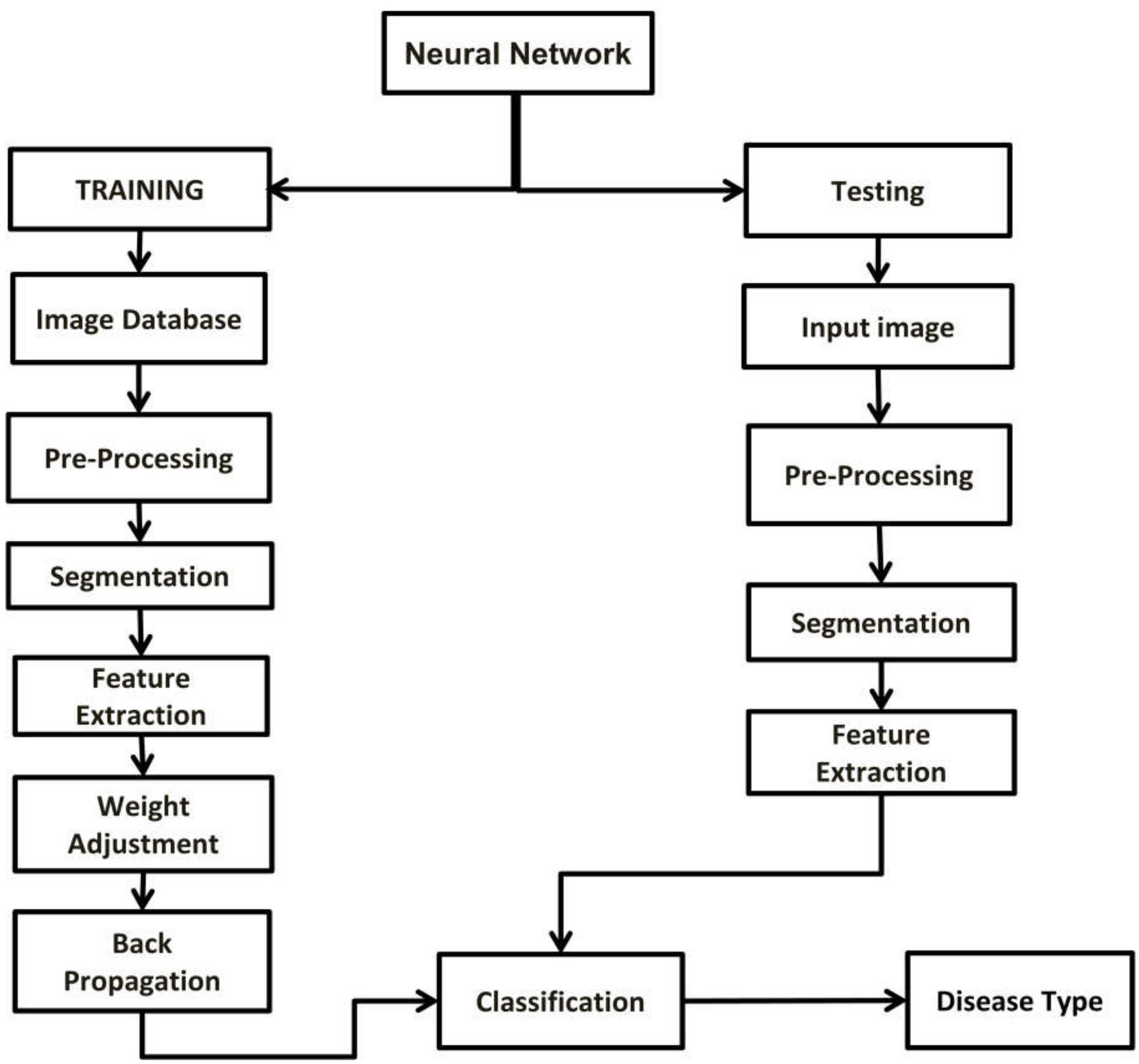

Fig 4: training and testing model using ANN

4. SIMULATION RESULTS:

For simulation purpose four different categories of plant leaf are considerer and trained using ANN through MATLAB R2018b simulation environment. Figure represents the training dataset with first row of images are belongs to Alternaria Alternata disease, second row of images are belongs to Anthracnose disease, third row of images are belongs to Bacterial Blight disease, fourth row of images are belongs to Cercospora Leaf Spot disease and finally fifth row of images are belongs to Healthy Leaves with no Healthy Leaves.

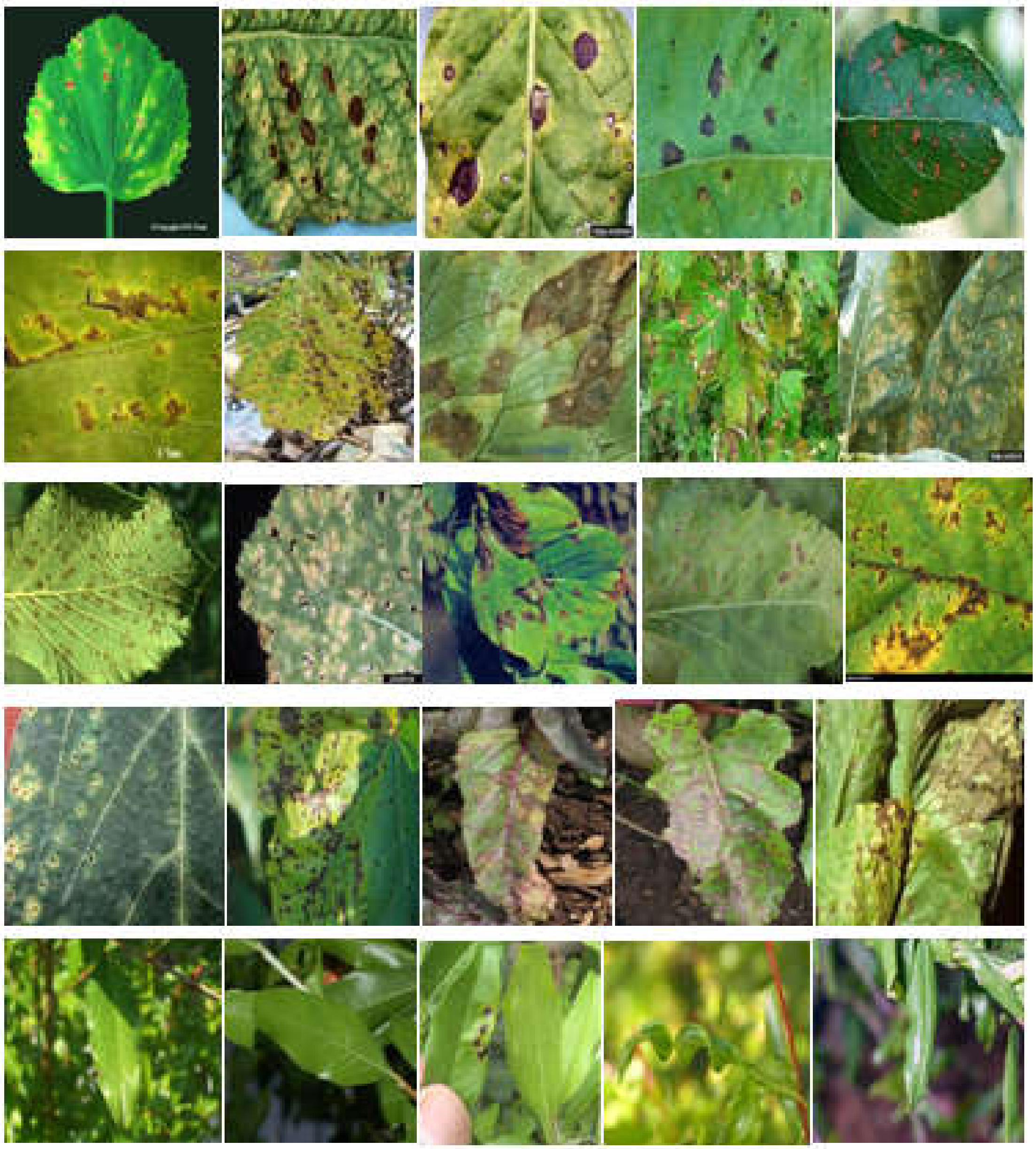

Fig 5: training dataset

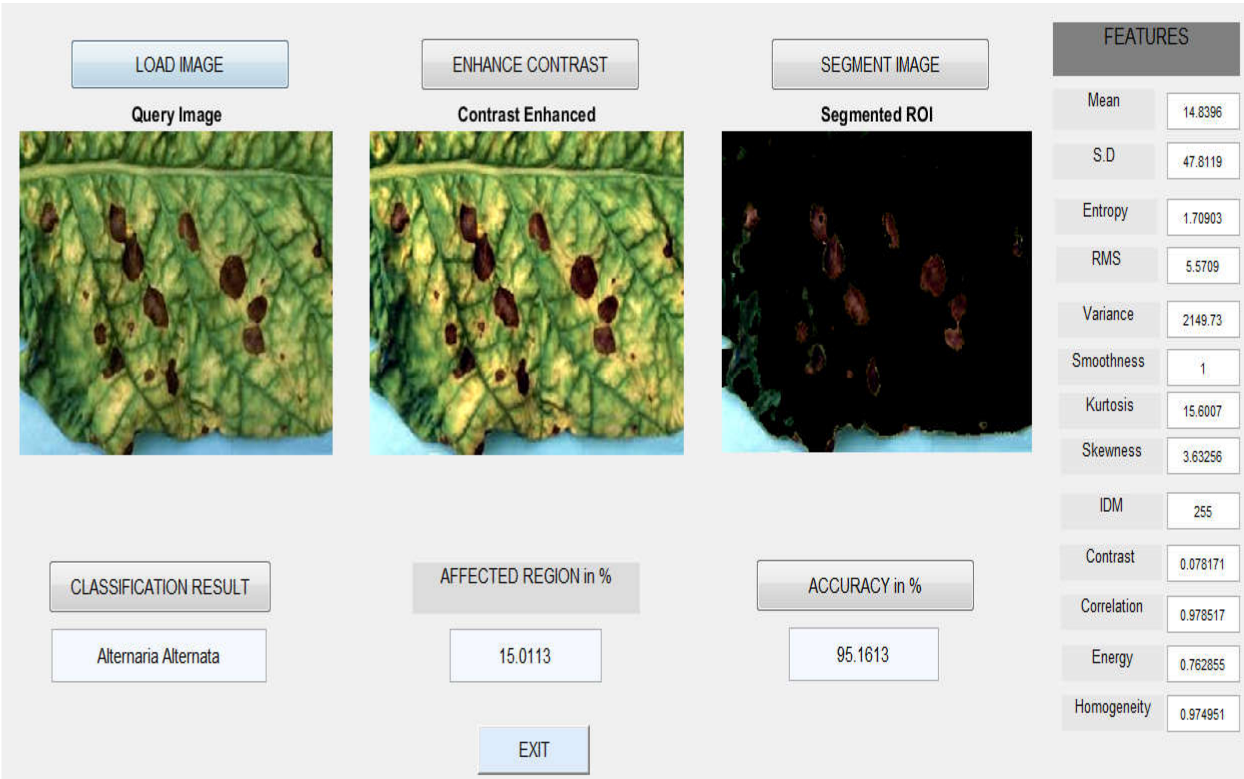

Fig 6: Alternaria Alternata disease detection and classification

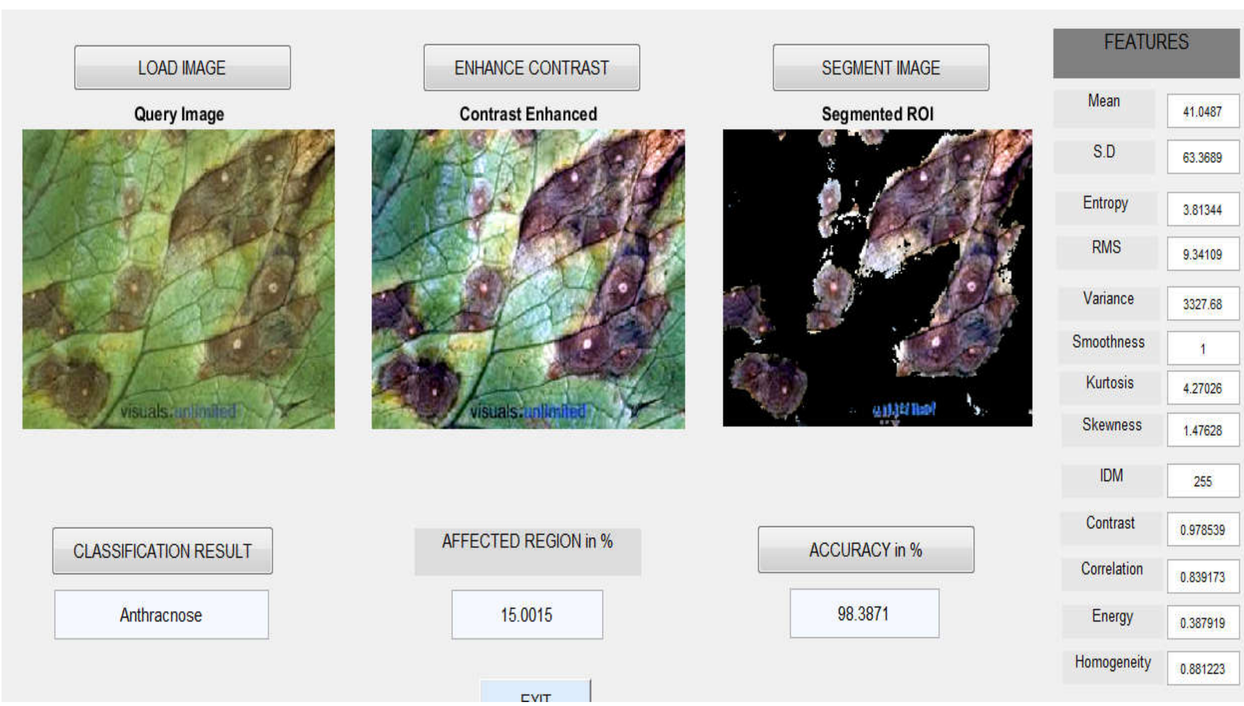

Fig 7: Anthracnose disease detection and classification

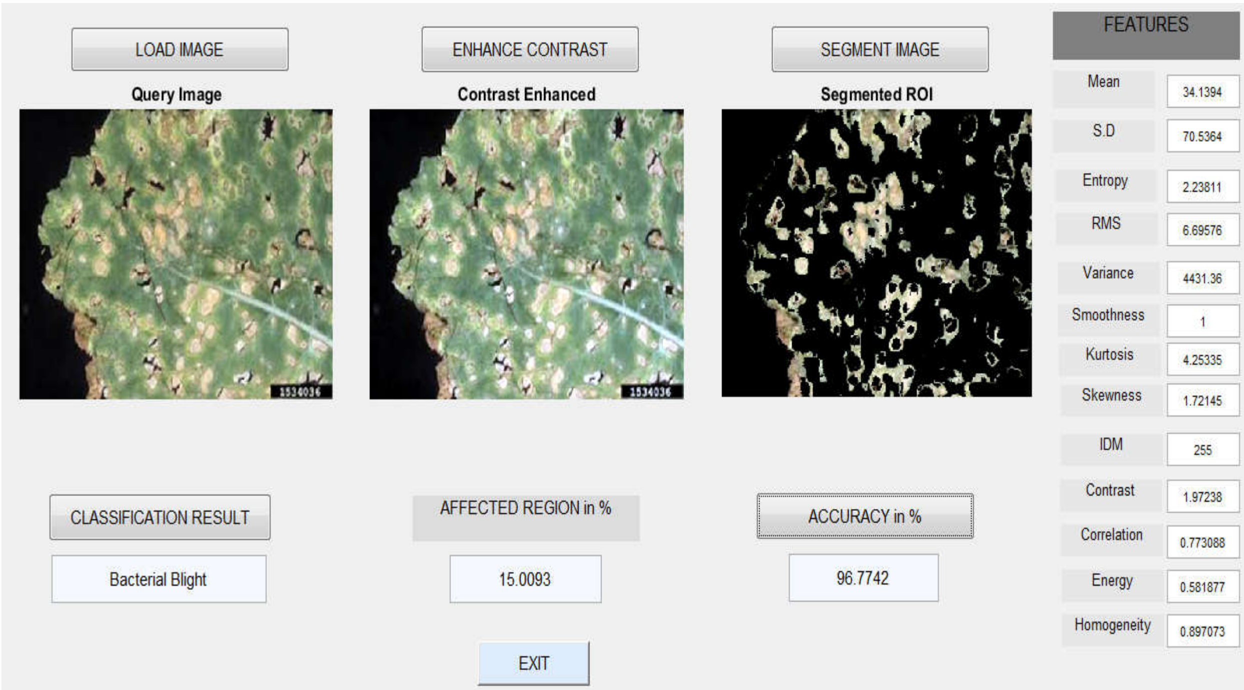

Fig 8: Bacterial Blight disease detection and classification

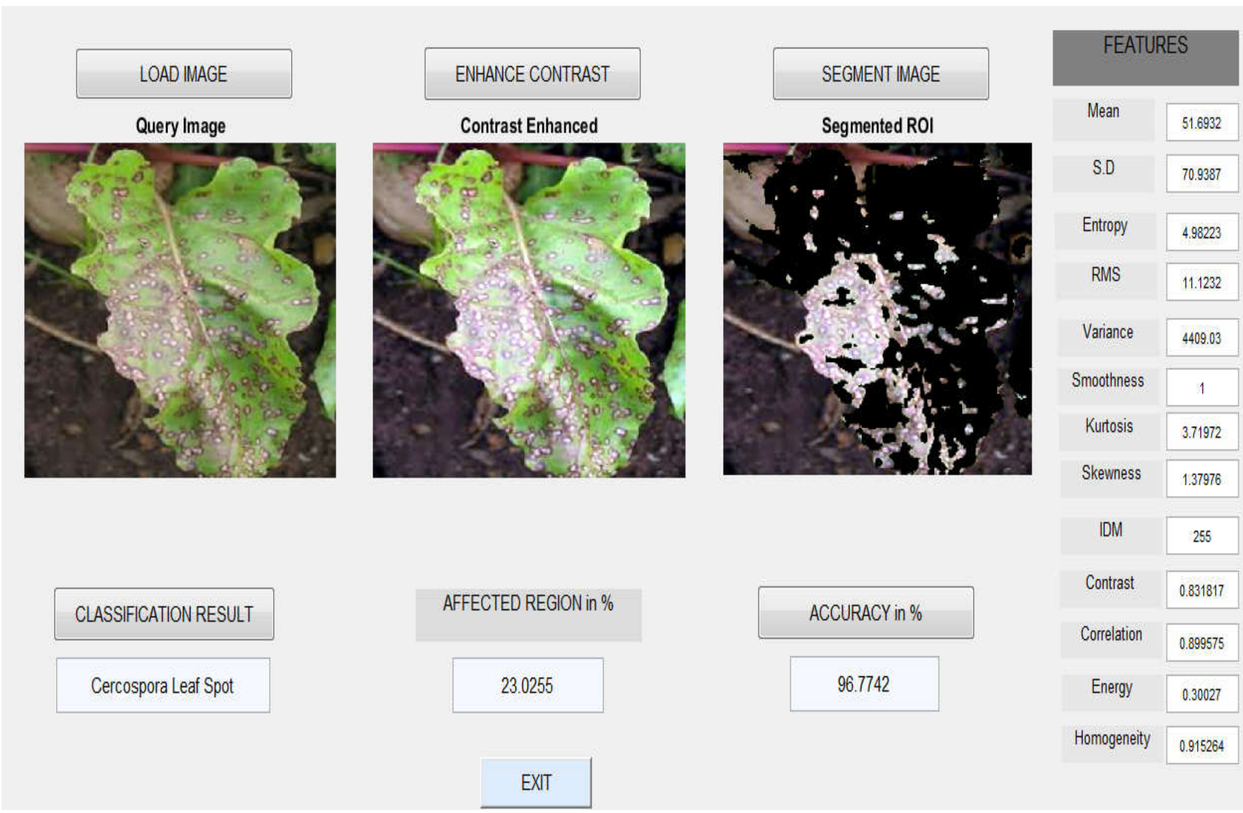

Fig 9: Cercospora Leaf Spot disease detection and classification

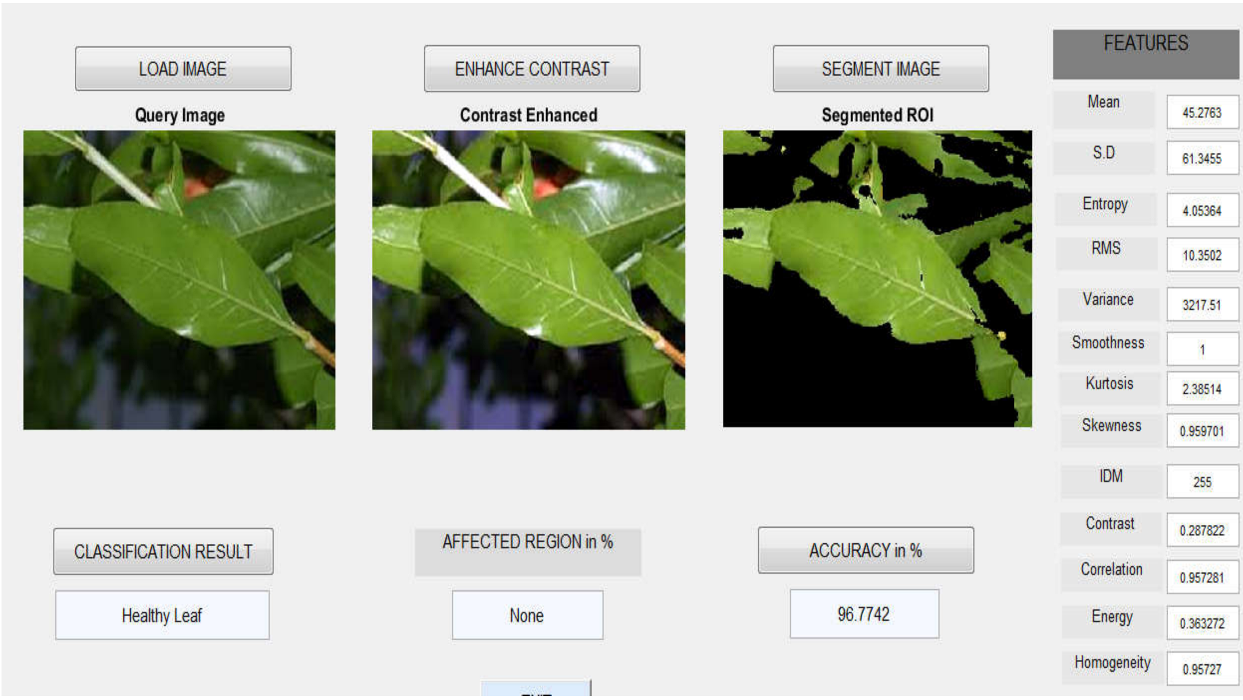

Fig 10: Healthy Leaf detection and classification

From figure 6 to 8 represents the four types of disease detection and classification results. And Fig 9 represents the detection and classification operation on healthy leaf image. In each figure query image represents the test input image, contrast enhanced image represents the pre-processed and noise removal of input image. Then segmented ROI represents the disease detected location in original input image. Then its features such as mean,SD,Entropy,RMS,Variance,Smoothness,kurtosis,skewness,IDM,contrast,correlation,energy and homogeneity are extracted using GLCM filter are presented as column in feature window. Finally recognised disease type is shown in Classification result, with its affected region and accuracy of classification. Average of all those features and parameters are summarized in table 1. The quantitative and qualitative results shows that our proposed method is performing much better in detection and classification compared to the existing SVM based classification.

Table 1: Summary of comparison

	SVM CLASSIFICATION[1]	ANN CLASSIFICATION
MEAN	0.01	0.03
STANDRAD DEVIATION	0.03	0.089
ENTROPY	2.14	3.17

RMS ERROR	1.15	0.089
VARIANCE	0.2054	0.0080
SMOOTHNESS	0.57	0.920
KURTOSIS	2.87	7.32
SKEWNESS	0.839	0.469
IDM	0.024	0.057
CONTRAST	0.149	0.2088
PSNR	25.28	41.20
ACCURACY	87%	96%
SENSITIVITY	67%	83%
SPECIFICITY	86%	93%

CONCLUSION: There are many methods in automated or computer vision plant disease detection and classification process, but still, this paper field is lacking. In addition, there are still no commercial solutions on the market, except those dealing with plant species recognition based on the leaves images. In this paper, a new approach of using deep learning method was explored in order to automatically classify and detect plant diseases from leaf images. The developed model was able to detect leaf presence and distinguish between healthy leaves and 4 different diseases, which can be visually diagnosed. The complete procedure was described, respectively, from collecting the images used for training and validation to image preprocessing and augmentation and finally the procedure of training the deep ANN and fine-tuning. Different tests were performed in order to check the performance of newly created model.

References

[1] Korkut UB, Göktürk ÖB, Yildiz O. Detection of plant diseases by machine learning. In2018 26th Signal Processing and Communications Applications Conference (SIU) 2018 May 2 (pp. 1-4). IEEE.

[2] Sandhu, Gurleen Kaur, and Rajbir Kaur. "Plant Disease Detection Techniques: A Review." 2019 International Conference on Automation, Computational and Technology Management (ICACTM). IEEE, 2019.

[3]Raghavendra, B. K. "Diseases Detection of Various Plant Leaf Using Image Processing Techniques: A Review." 2019 5th International Conference on Advanced Computing & Communication Systems (ICACCS). IEEE, 2019.

[4] Mishra, B., Nema, S., Lambert, M. and Nema, S., 2017, March. Recent technologies of leaf disease detection using image processing approach—A review. In 2017 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS) (pp. 1-5). IEEE.

[5] Kaur, Sukhvir, Shreelekha Pandey, and Shivani Goel. "Semi-automatic leaf disease detection and classification system for soybean culture." *IET Image Processing* 12, no. 6 (2018): 1038-1048.

[6] Sharath, D. M., Kumar, S. A., Rohan, M. G., & Prathap, C. (2019, April). Image based Plant Disease Detection in Pomegranate Plant for Bacterial Blight. In *2019 International Conference on Communication and Signal Processing (ICCSP)* (pp. 0645-0649). IEEE

[7]Annabel, L. Sherly Puspha, T. Annapoorani, and P. Deepalakshmi. "Machine Learning for Plant Leaf Disease Detection and Classification–A Review." *2019 International Conference on Communication and Signal Processing (ICCSP)*. IEEE, 2019.

[8]Maniyath, Shima Ramesh, et al. "Plant disease detection using machine learning." *2018 International Conference on Design Innovations for 3Cs Compute Communicate Control (ICDI3C)*. IEEE, 2018.

[9] Bharate, Anil A., and M. S. Shirdhonkar. "A review on plant disease detection using image processing." *2017 International Conference on Intelligent Sustainable Systems (ICISS)*. IEEE, 2017.

[10] Rashid, Mamoon. "Crop Disease Detection Using Image Processing Techniques-A Systematic Review." *Journal of the Gujarat Research Society* 21.10s (2019): 373-382.