Plant Leaf disease classification and detection with CNN and optimization

Anjana [1], Mr. Keshav Kishore [2]
[1, 2] *Department of Computer Science & Engineering, AP Goyal Shimla University, INDIA*

Abstract- The agricultural land mass is more than just being a feeding sourcing in today's world. Indian economy is highly dependent of agricultural productivity. Therefore in field of agriculture, detection of disease in plants plays an important role. To detect a plant disease in very initial stage, use of automatic disease detection technique is beneficial. For instance, a disease named little leaf disease is a hazardous disease. This is the one of the reasons that disease detection in plants plays an important role in agriculture field, as having disease in plants are quite natural. If proper care is not taken in this area then it causes serious effects on plants and due to which respective product quality, quantity or productivity is affected. This PROJECT presents an algorithm for image segmentation technique which is used for automatic detection and classification of plant leaf diseases. It also covers survey on different diseases classification techniques that can be used for plant leaf disease detection. Image segmentation, which is an important aspect for disease detection in plant leaf disease, is done by using SVM,ANN.

Keywords- *SVM, ANN, CNN, PSO*

I. INTRODUCTION

Image processing has been proved to be effective tool for analysis of images in various fields and applications. Agriculture Sector where the parameters like canopy, yield, quality of product were the important measures from the farmer's point of view. Many times the availability of expert and their services may consume a lot of time as well as expert advice may not be affordable. Image processing along with availability of communication network can change the situation of getting the expert advice well within time and at affordable cost since image processing was the effective tool for analysis of parameters [1]. This paper intends to focus on the survey of application of image processing in agriculture field such as imaging techniques, weed detection and fruit grading. The analysis of the parameters has proved to be accurate and less time consuming as compared to traditional methods [2]. Application of image processing can improve decision making for vegetation measurement, irrigation, fruit sorting, etc. Irrigation/Water stress occurs when the water supply to the plants was limited Fertilizers, pesticides and quality of yield were the major factors of concern in agriculture. Most of the time the expertise were required to analyze the problems and which may be time consuming and costlier issue in developing countries. Image processing was one of the tools which can be applied to measure the parameters (leaf area index (LAI), nitrogen (N) uptake, total chlorophyll (Chl) content) related to agronomy with accuracy and economy. Applications of image processing in agriculture can be broadly classified in two categories: first one Remote Sensing depends upon the imaging techniques and second one based on applications like Weed Detection [9].

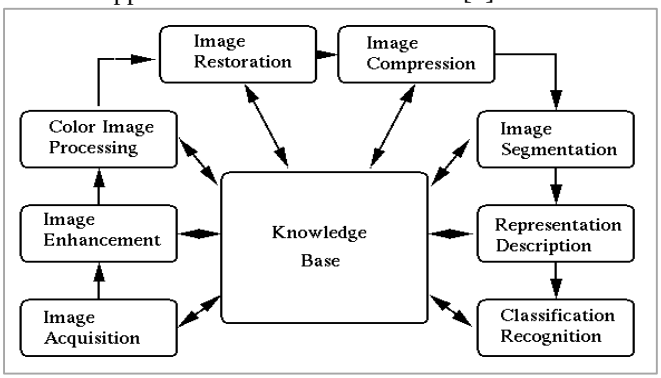

Fig.1 Steps Involved in Image Processing

A. Types of Image Processing

There are two types of methods used for image processing namely, analogue and digital image processing. Analogue image processing can be used for the hard copies like printouts and photographs. Image analysts use various fundamentals of interpretation while using these visual techniques. Digital image processing techniques help in manipulation of the digital images by using computers. The three general phases that all types of data have to undergo while using digital technique are pre-processing, enhancement, display images and information extraction.

- **Image Segmentation:** It is a process of dividing the images into the number of parts according to the parts that are strongly co-relate with the object in an image. It is done on the basis of discontinuity and similarity of the pixels.
- **Feature extraction:** Feature extraction is the process of transforming the input data into a set of features which can very well represent the input data. It is a special form of dimensionality reduction. In

agricultural applications features extracted on the basis of color and texture of crop and fruit.

- **Image classification:** It is process in which image is grouped together on the basis of distinct group or the parts that have common features.

B. Image processing in Agriculture

The application of image processing involves the following

- **Weed Detection:** Weed were the plants growing in wrong place in farm which compete with crop for water, light, nutrients and space which cause the loss in the crops.
- **Fruit Grading:** Due to increase in expectation of food safety and quality fruit grading is done on the basis of their color and weight. It is done by image processing very fast ant and effectively.
- **Monitoring of Drought:** By using the image processing method it is easy to found the drought areas.
- **Crop identification:** Used to identify which crop is sown and which is ready it is easy to identify by using image processing.
- **Crop condition and Stress detection:** Image processing is used to identify the crop condition after the rain or storms and also measure the Stress condition.

C. Leaf Disease Detection using Image Processing

Plant diseases are important factors, as it can cause significant reduction in both quality and quantity of crops in agriculture production. Therefore, detection and classification of diseases is an important task. Traditionally the naked eye observation of experts is the main approach adopted in practice for detection and identification of plant diseases. But, this requires continuous monitoring of experts which might be prohibitively expensive in large farms. Further, in some developing countries, farmers may have to go long distances to contact experts, this makes consulting experts too expensive and time consuming and moreover farmers are unaware of non-native diseases(belongs to other region) like weeds. Automatic detection of plant diseases is an important research topic as it may prove benefits in monitoring large fields of crops, and thus automatically detect the diseases from the symptoms that appear on the plant leaves [2]. This enables machine vision that is to provide image based automatic inspection, process control and robot guidance. Comparatively, visual identification is labor intensive, less accurate by Spatial Gray-level Dependence Method (SGDM). The texture features are calculated and the classification is done using squared distance technique.

Plant disease identification by visual way is more laborious task and at the same time, less accurate and can be done only in limited areas. Whereas if automatic detection technique is used it will take less efforts, less time and become more accurate. Image processing is used for measuring affected area of disease and to determine the difference in the color of the affected area. In plants, some general diseases seen are brown and yellow spots, early and late scorch, and others are fungal, viral and bacterial diseases [9].

- Red Rot Diseases in sugar cane
- Smut in Pearl Millet
- Tikka in Groundnut
- Blast in Rice
- Rust in Coffee
- Wilt in cotton

Fig.2 Diseases in Plants [17, 18, and 19]

D. Techniques Used for Leaf Disease Detection

1. K-mean Clustering: It is a supervised learning algorithm used to solve the clustering problems. It provides an easy way to classify the data into number of clusters. It assigns a group to each data point on the basis of their similarity. This algorithm provides the centroids of the K clusters which can be used to label new data. Labels are used for training data.

2. Neural Network: This networks works in the layer to provide the superior performance in image enhancement. In this method first layer is called as input layer, second layer is hidden layer and the last is output layer. Each layer performs a different function as their name. Types of neural network are following:-

- Back propagation Method
- Feed Forward Back Propagation Method Neural Network

3. Neuro-Fuzzy Inference System: This method is used as a classifier which classifies the texture of the images. It works on the basis of Fuzzy rules to provide the effective results. It gives efficient and optimal solution of the problem.

II. RELATED WORK

Arya et al. [1] proposed a plant diseases detection method by identifying the leaves using the concept of genetic algorithm

with image processing method. The diseases detection is an important part in the field of agriculture for this traditionally people observes by naked eyes but all time it is not possible to effectively classify the diseases this problem is solved by using the latest technologies. The proposed approach classifies the leaves from the images and then classifies them by effectively by using the genetic algorithm effectively. The classification is done on the basis of texture of the leaves. This approach helps in to identify the diseases in early stages.

Deisy et al. [2] introduced the image segmentation method for leaf diseases detection. This is done by feature extraction of the leaves images. In this work detection is based on the following methodology. Firstly the image acquisition after this preprocessing of images for noise removal from the images. The segmentation process is performed after preprocessing in segmentation image is divided into small parts and the feature extraction process is done. By using these features diseases detection is done. The result of the simulation proved that the proposed method is fast, effective and available at low cost. Joshi et al. [3] worked on the diseases detection in cotton plant. This work is based on the identification of cotton bug on cotton plant. The detection process is done by using hybrid fuzzy c-means method and thresholding method. The proposed approach divide the image into segment it enhance the accuracy of detection. The detection is done on the features like orientation, length and area. In this work neural network classifier is used for classification which enhances the accuracy in results. Zhang et al. [4] presented the method of leaf diseases detection by using the concept of K-mean clustering algorithm and pyramids of histogram method. In this work the image is divided into the super pixel clusters by using super pixel clustering algorithm. After this K-mean clustering algorithm is applied on the super pixel clusters. The PHOG features are extracted from the segmented image and then concatenate the 4- PHOG descriptor as a vector. The result of the proposed approach shows its effectiveness and proves that it works properly in diseases detection. Hossian et al [5] worked on the diseases detection and recognition on the Tea plant leaves. The detection process is done by using the support vector machine classifier. The detection process is based on the 11features for the images and later these features are used for the classification process. On the basis of image features diseases is analyzed and every time the image of leaves is uploaded into the SVM database. The uploaded image is matched with the images in the database for diseases recognition. The result of this process shows it takes less computation time with high accuracy and enhances the efficiency of detection and recognition. Shariff et al [6] presented disease detection and classification approach which is based on the weighted segmentation and feature selection. This work is based on the detection of citrus diseases in fruits. In this work firstly detection of lesion spot on the citrus fruit and leaves and after this classification of diseases is done. The lesion spots are extracted by using optimized weighted segmentation method. The effective features are selected by using the hybrid feature selection method which consist of entropy, PCA score and skewness-based covariance vector. After this process Multi class-SVM classifier is used for classification. This approach gives the high accuracy. Singh et al. [7] presented a review on the plant diseases detection techniques. This is done because diseases detection is an important part in the field of agriculture for this traditionally people observes by naked eyes but all time it is not possible to effectively classify the diseases this problem is solved by using the latest technologies. This review presented the different algorithms of machine learning and their working in diseases detection for effective accuracy. Khan et al. [8] proposed a diseases detection method by using the concept of multilevel segmentation and expectation maximization algorithm. In this work salient regions are extracted from the images by using binary partitioned tree. And it utilizes the principle of eigen vector. The accuracy of the proposed approach is higher than the existing approach. Picon et al. [9] worked on the diseases classification by using the concept of deep convolution neural network approach on the mobile captured image of the diseases. In this work firstly data set is trained and validated which contains the images of the diseases. This approach also analyzed the diseases in early, medium or late. After this mobile application is developed which detect the diseases. Liang et al. [10] proposed an approach for identify the soybean defoliation by using color image analysis. The leaf area is calculated by using distance classification method on the set of images. The edges of the leaf are detected by using Canny edge detection algorithm. The segmentation provides the effective performance in the classification of images. In this work two regression model is used that are polynomial and logistic regression.

Singh et al [11] proposed the leaf diseases detection by using the concept of the image segmentation and soft computing methods. The proposed approach automatically classifies and detects the diseases in leaf. The surveys of different diseases related to plants are also used in detection. This work is done with the help of Genetic algorithm which optimizes the variables effectively. Zhang, Shanwen, et al. [12] worked on the cucumber diseases detection by using the sparse representation classification. This work is done on the image of the leaf in which image is divided into small segments by using k-mean clustering. After these features of the image is extracted on the basis of shape and color features. The diseases classification is done by using the sparse matrix. The results of the proposed approach show that it reduces the computation time and improve the sparse representation. The result of this approach is also compared with the other approach. And it performs better than other.

Patil, et al. [13] presented the content based image retrieval for plant diseases detection on the basis of features like shape, color and texture features. The HSV color histogram method is used

to extract the color feature from the image. The shape features are provided by the Scale –Invariant feature transform. The texture feature like Local Binary Patterns and Gabor filters. In the proposed work Local and Gabor is analyzed. The proposed work is tested on the soybeans and at last the performance is enhanced by combining the shape, color and texture of features. Lu, Yang, et al. [14] presented the deep neural network for the identification of the rice diseases. In this work CNN is trained to identify the 10 rice diseases. The three CNN layer are used to perform the actions in which first layer extract the features from image like edges, lines and corners. The other CNN layer gets high level features. After this pooling layer perform the action and them softmax layer classify the diseases. The simulation result of this approach is compared with other approaches and it gives effective accuracy rate and the false report rate is 0.

III. THE PROPOSED METHOD

In this section, we discussed the proposed approach and the methodology used to achieve the results.

A. Proposed Technique: The proposed technique involves the following steps:

Step1: Input the image.
Step2: Preprocessing of the image which reduces the noisy data from it.
Step3: Image segmentation is performed which divides the image into the small segments.
Step4: Extract the features from the segments of the image.
Step5: Selects the optimized features by optimization process using Deep learning method.
Step6: Features learned by the classifier
Step7: Detect the affected leaf
Step8: Analysis of Accuracy, Precision, and recall.

B. Proposed methodology: Algorithm/Flowchart

C. Algorithm

Following are the algorithms that are used in the proposed work. The Grey wolf optimization are used to optimize the features which are given by CNN.

(a) CNN: Neural Network and Deep Learning How to find effective features is the core issue in image classification and pattern recognition. Humans have an amazing skill in extracting meaningful features, and a lot of research projects have been undertaken to build an FE system as smart as human in the last several decades. Deep learning is a newly developed approach aiming for artificial intelligence. Deep learning-based methods build a network with several layers, typically deeper than three layers. Deep neural network (DNN) can represent complicated data. However, it is very difficult to train the network. Due to the lack of a proper training algorithm, it was difficult to harness this powerful model until Hinton and Salakhutdinov proposed a deep learning idea [27]. Deep learning involves a class of models that try to learn multiple levels of data representation, which helps to take advantage of input data such as image, speech, and text. Deep learning model is usually initialized via unsupervised learning and followed by fine-tuning in a supervised manner. The high level features can be learnt from the low-level features. This kind of learning leads to the extraction of abstract and invariant features, which is beneficial for a wide variety of tasks such as classification and target detection. There are a few deep learning models in the literature, including DBN [14], [15] SAE and CNN. Recently, CNNs have been found to be a good alternative to other deep learning models in classification and detection. In this paper, we investigate the application of deep CNN for HSI FE. The architecture of CNN is different from other deep learning models. There are two special aspects in the architecture of CNN, i.e., local connections and shared weights. CNN exploits the local correlation using local connectivity between the neurons of near layers. We illustrate this in Fig.4, where the neurons in the mth layer

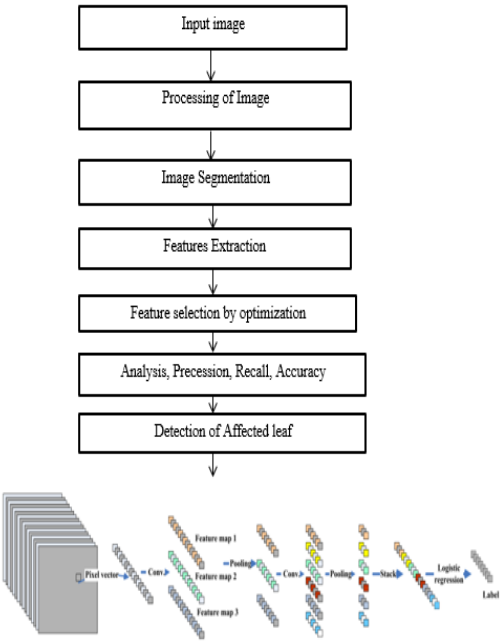

Fig.3 Proposed Flowchart

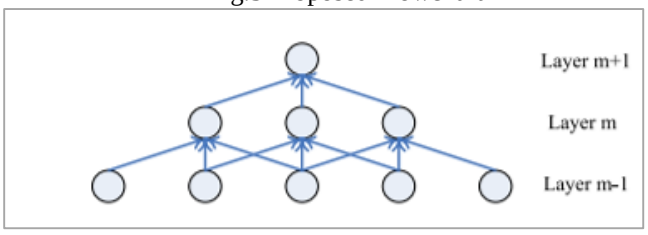

Fig.4 Convolution Neural Network [14]

(b) Grey Wolf Optimization

Grey Wolf optimization algorithm is a bio-inspired algorithm which is based on the leadership and hunting behavior of the wolves in the pack. The grey wolves prefer to live in the pack which is a group of approximate 5-12 wolves. In the pack each member has social dominant and consisting according to four different levels. The below given figure shows the social hierarchy of the wolves which plays and important role in hunting

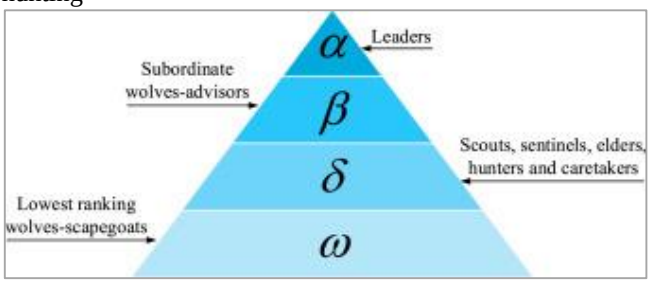

Fig.5 GWO Hierarchy

1. The wolves on the first level are called alpha wolves (α) and they are leaders in the hierarchy. Wolves at this level are the guides to the hunting process in which other wolves seek, follow and hunt and work as a team. Decision making is the main task that is performed by the alpha wolves and the order by the alpha wolves is followed by all members of the pack.
2. Second level wolves are called beta (β). These wolves are called subordinates and advisors of alpha nodes. The beta wolf council helps in decision making. Beta wolves transmit alpha control to the entire packet and transmit the return to alpha.
3. The wolves of the third level are called Delta wolves (δ) and called scouts. Scout wolves at this level are responsible for monitoring boundaries and territory. The sentinel wolves are responsible for protecting the pack and the guards are responsible for the care of the wounded and injured.
4. The last and fourth level of the hierarchy are called Omega (ω). They are also called scapegoats and they must submit to all the other dominant wolves. These wolves follow the other three wolves.

IV. RESULT ANALYSIS

In Experiment divided on two parts

1. Classification of Disease and non-disease leaf
2. Detection of particular leaf as disease or not disease

A. Parameters

Accuracy: Accuracy is defined by the ratio of correct prediction over the total number of cases to be evaluated.

$$\text{Accuracy}=\frac{\text{True Positives}+\text{True Negatives}}{\text{True Positives}+\text{False Postives}+\text{True Negatives}+\text{False Negatives}}$$

Precision: Precision defined by the ratio of true positives number (T_p) over true positives number (T_p) plus false positives number (F_p)

$$\text{Precision}=\frac{\text{True Positives}}{\text{True Positives}+\text{Fasle positives}}$$

Recall: The algorithm performance is measured through recall which is true positives rate (TPR) equivalent.

$$Recall=\frac{\text{True positives}}{\text{True positives}+\text{False Negatives}}$$

F-measure: Root means squared method is used to calculate the RMS error.

$$\text{F-measure}=\sqrt{(\text{True negatives}-\text{True Positives})^2+(\text{False Negatives}-\text{Fasle Positives})^2}$$

5.2 Classification Results

Approaches	CNN	CNN_PSO	Neural Network	SVM
Accuracy	96.7	98.45	76.56	46.73

Table.1 Accuracy comparison of different classifiers

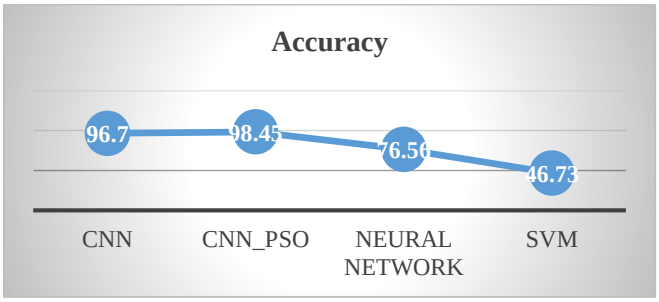

Fig.6 Comparison of accuracy of different classifier

In fig.6 comparison of convolution neural network and convolution neural network with particle swarm optimization and also SVM and ANN. In fig.6 X-axis classifier and y-axis value of accuracy which show that CNN_PSO show 98.45% significant accuracy and CNN 96.7% these accuracies highly significant from SVM and ANN. These accuracy increase in case of CNN because convolution the features and non-linear mapping so effective features given effective accuracy but in CNN_PSO optimize features and more improve CNN.

Table.2 Precision comparison of different classifiers

Approaches	CNN	CNN_PSO	Neural Network	SVM
Precision	90.34	94.56	74	43.73

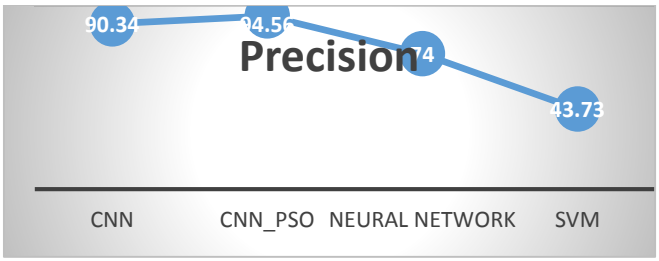

Fig.7 Comparison of precision of different classifier

In fig.7 comparison of convolution neural network and convolution neural network with particle swarm optimization and also SVM and ANN. In fig5.1 X-axis classifier and y-axis value of Precision which show that CNN_PSO show 98.45% significant precision and CNN 96.7% these accuracies highly significant from SVM and ANN. These precision increase in case of CNN because convolution the features and non-linear mapping so effective features given effective precision but in CNN_PSO optimize features and more improve CNN.

Table.3 Recall comparison of different classifiers

Approaches	CNN	CNN_PSO	Neural Network	SVM
Recall	92.14	96.78	75.45	42.73

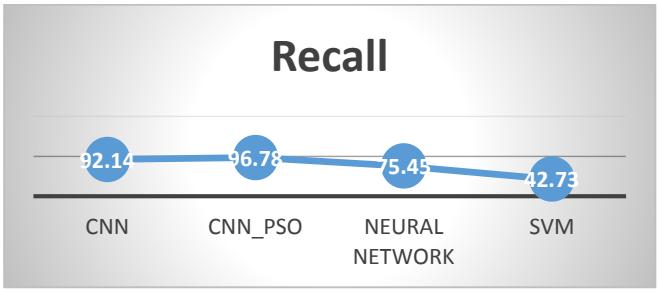

Fig.8 Comparison of Recall of different classifier

In fig.8 comparison of convolution neural network and convolution neural network with particle swarm optimization and also SVM and ANN. In fig.8 X-axis classifier and y-axis value of recall which show that CNN_PSO show 98.45% significant recall and CNN 96.7% these recall highly significant from SVM and ANN. These recall increase in case of CNN because convolution the features and non-linear mapping so effective features given effective recall but in CNN_PSO optimize features and more improve CNN.

Table.4 F-measure comparison of different classifiers

Approaches	CNN	CNN_PSO	Neural Network	SVM
F-measure	93.7	97.67	76.56	41.73

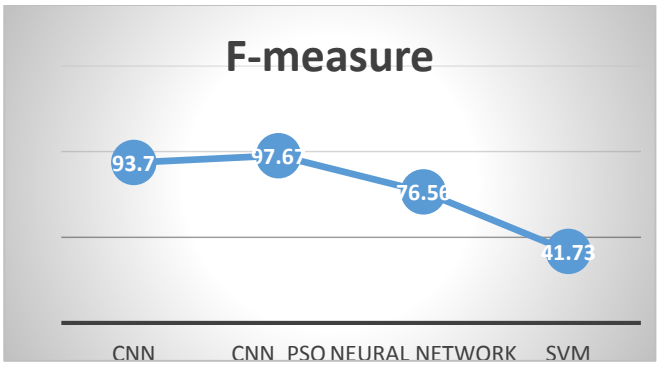

Fig.9 Comparison of F-measure of different classifier

In fig.9 comparison of convolution neural network and convolution neural network with particle swarm optimization and also SVM and ANN. In fig.9 X-axis classifier and y-axis value of F-measures which show that CNN_PSO show 98.45% significant F-measures and CNN 96.7% these F-measures highly significant from SVM and ANN. These F-measures increase in case of CNN because convolution the features and non-linear mapping so effective features given effective F-measures but in CNN_PSO optimize features and more improve CNN.

Table.5 comparison of different classifiers

Approaches	Accuracy	Precision	Recall	F-measure
CNN	96.7	90.34	92.14	93.7
CNN_PSO	98.45	94.56	96.78	97.67
Neural Network	76.56	74	75.45	76.56
SVM	46.73	43.73	42.73	41.73

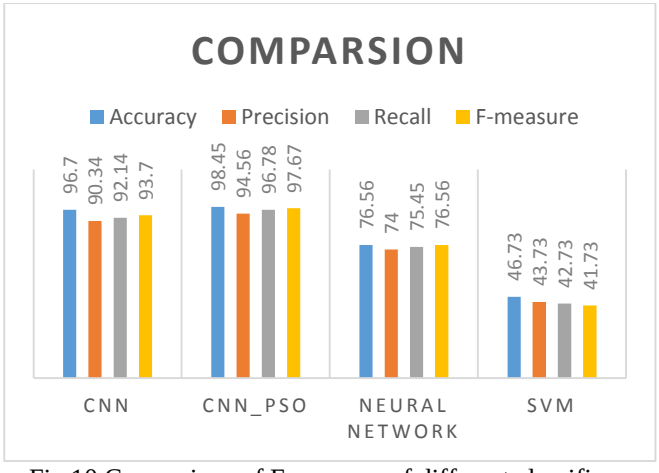

Fig.10 Comparison of F-measure of different classifier

5.3 Detection Screen Shot

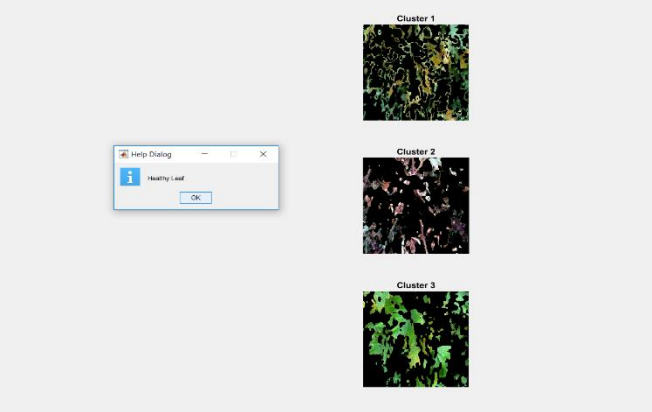

Fig.11 Detection leaf-disease by CNN_PSO

In fig.11 process of detection leas has disease or not disease and if disease which type otherwise it will indicate normal. This detection method uses from above given results. In analysis of detection use CNN_PSO. In fig.11 show three cluster which indicate different layer segmentation and features which detect neural network and show healthy leaf.

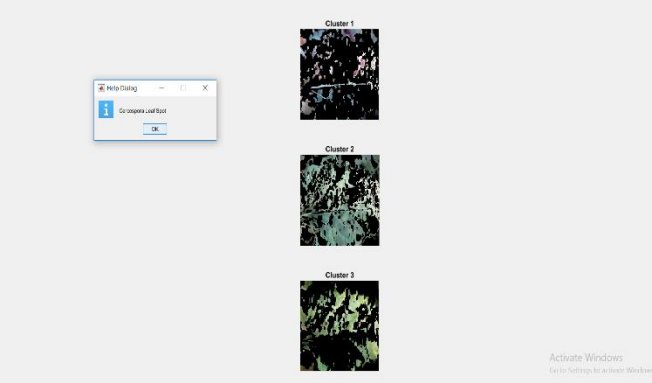

Fig.12 Detection leaf disease by CNN_PSO

In fig.12 process of detection leas has disease or not disease and if disease which type otherwise it will indicate normal. This detection method uses from above given results. In analysis of detection use CNN_PSO. In fig.11 show three cluster which indicate different layer segmentation and features which detect neural network and show comosporasa leaf disease.

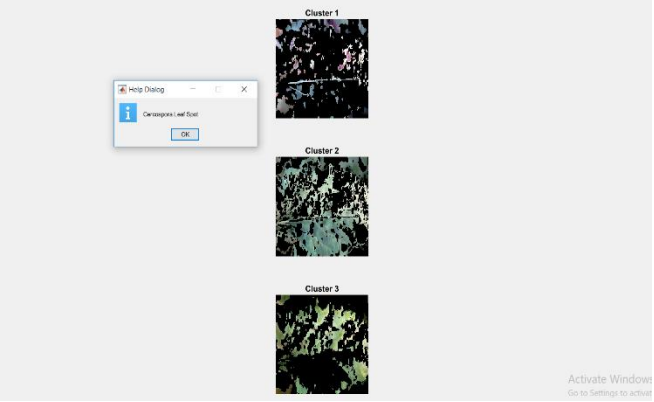

Fig.13 Detection leaf disease by CNN_PSO

In fig.13 process of detection leas has disease or not disease and if disease which type otherwise it will indicate normal. This detection method uses from above given results. In analysis of detection use CNN_PSO. In fig.11 show three cluster which indicate different layer segmentation and features which detect neural network and show corospora leaf disease.

V. CONCLUSION

The existing method for plant disease detection is simply naked eye observation by experts through which identification and detection of plant diseases is done. For doing so, a large team of experts as well as continuous monitoring of plant is required, which costs very high when we do with large farms. At the same time, in some countries, farmers do not have proper facilities or even idea that they can contact to experts. Due to which consulting experts even cost high as well as time consuming too. In such conditions, the suggested technique proves to be beneficial in monitoring large fields of crops. Automatic detection of the diseases by just seeing the symptoms on the plant leaves makes it easier as well as cheape rn different diseases classification techniques used for plant leaf disease detection and an algorithm for image segmentation technique that can be used for automatic detection as well as classification of plant leaf diseases later. Therefore, related diseases for these plants were taken for identification. With very less computational efforts the optimum results were obtained, which also shows the efficiency of proposed algorithm in recognition and classification of the leaf diseases. Another advantage of using this method is that the plant diseases can be identified at early stage or the initial stage. To improve recognition rate in classification process Artificial Neural Network, SVM.

VI. REFERENCES

[1] Arya, M. S., K. Anjali, and Divya Unni. "Detection of unhealthy plant leaves using image processing and genetic algorithm with Arduino." *2018 International Conference on Power, Signals, Control and Computation (EPSCICON)*. IEEE, 2018.

[2] Deisy, C., and Mercelin Francis. "Image Segmentation for Feature Extraction: A Study on Disease Diagnosis in Agricultural Plants." *Feature Dimension Reduction for Content-Based Image Identification*. IGI Global, 2018. 232-257.

[3] Joshi, Ms KV, and D. D. Shah. "Hybrid of the Fuzzy C Means and the Thresholding Method to Segment the Image in Identification of Cotton Bug." *International Journal of Applied Engineering Research* 13.10 (2018): 7466-7471.

[4] Zhang, Shanwen, et al. "Plant diseased leaf segmentation and recognition by fusion of superpixel, K-means and PHOG." *Optik-International Journal for Light and Electron Optics* 157 (2018): 866-872.

[5] Hossain, Md Selim, et al. "Recognition and detection of tea leaf's diseases using support vector machine." *Signal Processing & Its Applications (CSPA), 2018 IEEE 14th International Colloquium on*. IEEE, 2018.
[6] Sharif, Muhammad, et al. "Detection and classification of citrus diseases in agriculture based on optimized weighted segmentation and feature selection." *Computers and Electronics in Agriculture* 150 (2018): 220-234.
[7] Singh, Jaskaran, and Harpreet Kaur. "A review on: Various techniques of plant leaf disease detection." *2018 2nd International Conference on Inventive Systems and Control (ICISC)*. IEEE, 2018.
[8] Khan, Zia Ullah, et al. "Automatic detection of plant diseases; utilizing an unsupervised cascaded design." *Applied Sciences and Technology (IBCAST), 2018 15th International Bhurban Conference on*. IEEE, 2018.
[8] Picon, Artzai, et al. "Deep convolutional neural networks for mobile capture device-based crop disease classification in the wild." *Computers and Electronics in Agriculture* (2018).
[9] Liang, Wei-zhen, Kendall R. Kirk, and Jeremy K. Greene. "Estimation of soybean leaf area, edge, and defoliation using color image analysis." *Computers and Electronics in Agriculture* 150 (2018): 41-51.
[10] Singh, Vijai, and A. K. Misra. "Detection of plant leaf diseases using image segmentation and soft computing techniques." *Information Processing in Agriculture* 4.1 (2017): 41-49.
[11] Zhang, Shanwen, et al. "Leaf image based cucumber disease recognition using sparse representation classification." *Computers and electronics in agriculture* 134 (2017): 135-141.
[12] Patil, Jayamala Kumar, and Raj Kumar. "Analysis of content based image retrieval for plant leaf diseases using color, shape and texture features." *Engineering in Agriculture, Environment and Food* 10.2 (2017): 69-78.
[13] Lu, Yang, et al. "Identification of rice diseases using deep convolutional neural networks." *Neurocomputing* 267 (2017): 378-384.
[14] G. E. Hinton, S. Osindero, and Y. Teh, “A fast learning algorithm for deep belief nets,” Neural Comput., vol. 18, no. 7, pp. 1527–1554, Jul. 2006.
[15] N. LeRoux and Y. Bengio, “Deep belief networks are compact universal approximators,” Neural Comput., vol. 22, no. 8, pp. 2192–2207, Aug. 2010.