Review on Plant Leaf disease classification and detection with machine learning and optimization

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

Abstract- 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 and 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

Keywords—*image processing, leaf, optimization, machine learning*

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, 2]. 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 [5, 9].

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 [10]. 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.

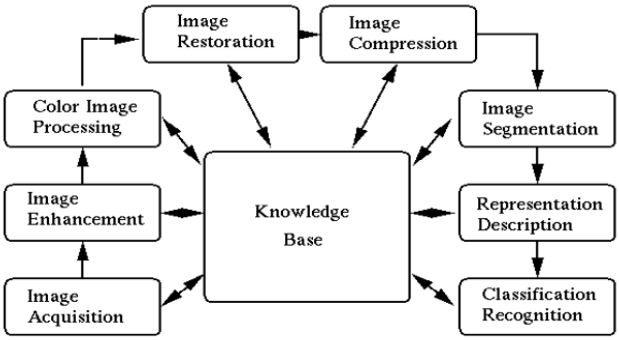

Fig.1 Steps Involved in Image Processing

- *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 [7]. 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. Applications of Image processing in Agriculture

- *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 [8]. 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.

II. 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 [11, 13]. 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, 12]. 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, 15].

- 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 [13, 14, and 15]

A. Techniques Used for Leaf Disease Detection

Some of the techniques have been discussed below:

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
- 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.

III. 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

Table.1 Inferences Drawn

Author's Name	Year	Algorithm Used/ Technology Used	Outcomes
Arya, et al.	2018	Genetic Algorithm	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.
Deisy et al.	2018	Segmentation and feature extraction	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.
Joshi et al	2018	Fuzzy C Mean	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.
Zhang et al	2018	Segmentation and Recognition	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.
Hossain et al.	2018	Support Vector Machine	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.
Sharif, et	2018	Segmentation and	This work is based on the detection of citrus diseases in fruits. In this work

al.		Feature Selection	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.
Singh, et al.	2018	Review Different detection methods	This review presented the different algorithms of machine learning and their working in diseases detection for effective accuracy.
Khan, et al.	2018	Binary Partitioned Tree	In this work salient regions are extracted from the images by using binary partitioned tree. And it utilizes the principle of eigen vector.

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. Zhang, Shanwen, et al. [9] 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. [10] 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. [11] 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. Bai, Xuebing, et al. [12] proposed fuzzy c-mean algorithm for identifying the cucumber diseases from leaf. The half target is isolated by using the marked watershed algorithm on the basis of HIS space. After this the distance is calculated between the pixel and cluster heads. The neighborhood mean gray value is calculated on the pixel which contains the 2-D vector and grey information. After this these
values are weighted by matrix w. The results show the efficiency and effectiveness of the proposed work by reducing the running time.

III. ALGORITHM

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.

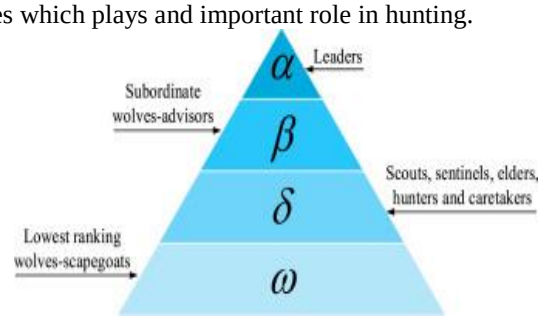

Fig.3 GWO Hierarchy

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.

V. 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.
[9] Zhang, Shanwen, et al. "Leaf image based cucumber disease recognition using sparse representation classification." *Computers and electronics in agriculture* 134 (2017): 135-141.
[10] 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.
[11] Lu, Yang, et al. "Identification of rice diseases using deep convolutional neural networks." *Neurocomputing* 267 (2017): 378-384.
[12] Bai, Xuebing, and et al. "A fuzzy clustering segmentation method based on neighborhood grayscale information for defining cucumber leaf spot disease images." *Computers and Electronics in Agriculture* 136 (2017): 157-165.
[13] Prajapati, Bhumika S., Vipul K. Dabhi, and Harshadkumar B. Prajapati. "A survey on detection and classification of cotton leaf diseases." *Electrical, Electronics, and Optimization Techniques (ICEEOT), International Conference on*. IEEE, 2016.
[14] Tejonidhi, M. R., et al. "Plant disease analysis using histogram matching based on Bhattacharya's distance calculation." *Electrical, Electronics, and Optimization Techniques (ICEEOT), International Conference on*. IEEE, 2016.
[15] Pinto, Loyce Selwyn, et al. "Crop disease classification using texture analysis." *Recent Trends in Electronics, Information & Communication Technology (RTEICT), IEEE International Conference on*. IEEE, 2016.