Detection and Classification of Orange Leaf Disease using Svm and K-Means Algorithm

Ravinderjit Kaur[1], Amardeep Kaur[2]
[12]*Punjabi University Regional Centre for Information Technology and Management , Phase 7, SAS Nagar (Mohali)*

Abstract- With the development of technology, different chemicals have been used on plants. Moreover, plants are easily affected by the diseases due to the different factors such as environmental conditions, an economic issue, fungal infection and so forth. It may affect the lives of social beings. Hence, it is necessary to detect and treat the disease of plant at correct time. Generally, detection area of plant disease is leaf. The disease affects the excellence and capacity of the crops. The detection of leaf disease takes maximum time. Therefore, an automated technique is used for detection of the leaf disease. The detection and the classification of the leaf disease avoid the loss in agricultural science. The condition of the plant leaf must be good to retain the safety and sustainability of the crops. However, detection of the plant disease is necessary to improve the productivity of the plant leaf. Detection of the plant leaf can be determined on the basis of the color, strength, area of attack, dimension and so forth. In existing research, the plant disease was identified and detected. The main goal of the existing research was the detection of the Psidium guajava (guava) leaf through digital image processing using K-means and SVM classification technique. In the proposed research, detection and classification of the orange leaf disease achieved from database. The images are acquired through pre-processing using various image processing methods like as area-based, edge, and model-based. After this process, image enhancement is done through boundary or area detection and spot detection algorithm. In addition, segmentation of the image is done by using k-means clustering method. K-means Clustering method used to divide the image into different types of clusters (Final cluster, Cluster 1 and Cluster 2). It defines the number of index, clusters, distance calculation based on maximum and minimum values. The Clustered information applied the HOG method to extract the global features in the form of feature vectors. It found the genuine features with orientation gradient method. After that, it has analyzed and detected using machine learning method to classify the disease in the orange leaves. The experimental analysis using MATLAB 2016a tool in GUI (Graphical User Interface) designed. Performance evaluation with Accuracy Rate, Precision, recall and compared with the existing methods.

Keywords- Orange Leave Detection, Classification (SVM), Clustering (K-means) and HOG feature extraction method.

I. INTRODUCTION

Agriculture is an essential source for humans as well as animals. The main requirement of the social beings is in the form of the nutrition which may be fruits, vegetables, puffed rice and crops[1]. It was investigated that about 10-20% loss were found in harvesting of the crops every year [2][3]. According to nutrition and farming department, 10 billion of population will be expected to hit by 2050. Hence, construction of the farming is required to increase up to 75% to facilitate the developed inhabitants and residents [4]. In contrast, maximum usage of the compounds like as disinfectants, and nematicides for controlling the infection on plant and leaves may lead to adversative effect on environmental situations [5][6]. Various kinds of the diseases among plants and leaves may affect trade and industry, community and environmental factors. Hence, detection and diagnosis of the disease become an essential approach [7] [8]. Some of the harmful effects of the plant disease are[9],

(i)Photosynthetic effect on plant leaves.

(ii)Influence progression and fabrication of plants.

(iii) Great impact on the growth and production of the country.

Among consideration of the some issues in plants, monitoring and detection of the plants is an important. Some of the detection and identification are methods are described as [10] [11],

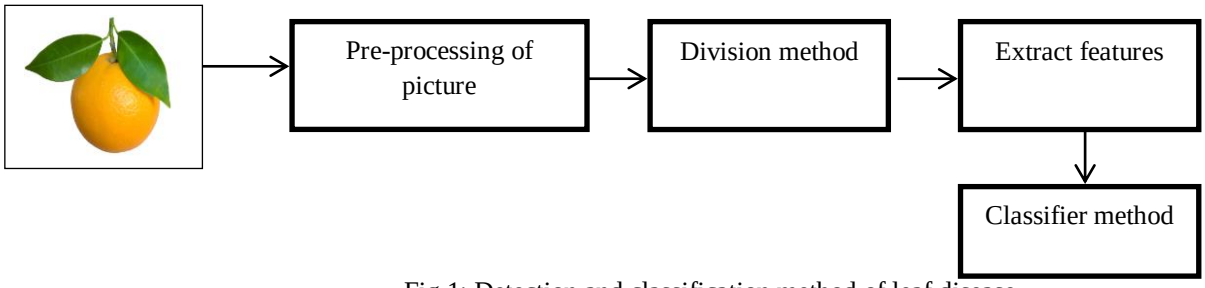


Fig.1: Detection and classification method of leaf disease

(i) Picture Acquisition: Digital pictures are acquired from digital camera and that is taken as the input picture. The leaf disease is required to be identified by computerised scheme.

(ii) Image Pre-processing: In this method, size of the picture can be changed, noise is removed and smoothening and enhancement of picture cab be done. The main stages includes in image pre-processing are [12]:

a) Extending compactness: Various methods are designed using specific picture where pictures are difficult to interpret.

b) Filtration of the noise: This method is utilised to extract different types of the distortion like as salt and pepper, periodical and Gaussian noise.

c) Adaptation of histogram: This is the method to improve the characteristics of picture.

(iii) Segmentation: In this technique, picture is segmented into various areas which consist of the pixels having similar features. The division of picture in segmentation method like as threshold, cluster based and water-shed technique [13][14].

(iv) Classify the disease: The disease is categorised as already defined database. The models utilised for the classification are support vector machine, multiple class SVM, and neural network[15].

In existing research, detection and identification of plant disease was done by image processing method. The main objective of the research was identification of Psidium guajava (guava) leaf using digital picture processing along with k-mean and SVM method. Segmentation of the leaf image was done through k-mean and classified through machine learning approach. It was determined that there is gradual reduction in the construction of the guava.

In the proposed research, detection and classification of the orange leaf disease achieved from database. The images are acquired through pre-processing using various image processing methods like as area based, edge and model-based. After this process, image enhancement is done through boundary or area detection and spot detection algorithm. In addition, segmentation of the image is done by using k-mean clustering method. K-means Clustering method used to divide the image into different types of clusters (Final cluster, Cluster 1 and Cluster 2). It defines the number of index, clusters, distance calculation based on maximum and minimum values. Clustered information applied the HOG method to extract the global features in the form of feature vectors. It found the genuine features with orientation gradient method. After that, it has analysed and detected using machine learning method to classify the disease in the orange leaves. The experimental analysis using MATLAB 2016a tool in GUI (Graphical User Interface) designed. Performance evaluation with Accuracy Rate, Precision, recall and compared with the existing methods.

II. LITERATURE SURVEY

Sunny, S. and Gandhi, M. I. et al., 2018[16] aimed for identification and classification of the disease of canker in orange leaf through image processing method. The proposed research based on two schemas for improving the clearness of the leaf picture. The main step used compact partial adapted histogram equalization (CLAHE) at the initial process. This enhanced the divergence of the disease and distributes the ROI utilizing k mean cluster approach and extract features of text through numerical GLCM. In next procedure, the classification of canker leaf picture was done through support vector machine. It improved the technique in lemon citrus disease recognition. The experimental analysis was done to improve the accuracy and decrease the time consumption during recognition of the canker disease. **Singh, V., and Misra, A. K. et al., 2017[17**] proposed research on detection and classification of the signs of plant disease using image segmentation process. The research was done on different methods of the detection of plant disease through genetic algorithm. This proposed approach was tested using various species of the plant. The exact outcome was acquired to get the exact detection and classification of the plant disease. The main goal of the algorithm was to detect the plant disease at an early stage. **Barbedo, J. G. A. et al., 2013[18]**presented a research on techniques based on digital IP method for detection and classification of the leaf disease from digital pictures. The observed leaves and stems were measured to determine the signs of the leaf disease. The method depends on two motives which were related to origins and seeds of the leaf. The determined technique was reliant on components which were described as, detect, quantify and classify. This research was based on the pathological factors of the leaf and pattern analysis. **Ashourloo, D., Aghighi, H., Matkan, A. A., Mobasheri, M. R and Rad, A. M. et al., 2016[19]** proposed research on measurement of signs of the diseased and non-infected leaves through non image spectra diameter in electric magnet area ranges from 360 to 2600 nm. The productions of the exact database were assigned through images from digital cameras to calculate the signs of disease. The modified samples of the collective database were used to sequence of each phase. Root mean square error was utilized to determine the scale level of the leaf. In addition, the signs of the various diseases of the disease were determined through machine learning method. **Soni, P. and Chahar, R. et al., 2016[20]**established an improvedvigorous probable neural network method for recognition of disease of various leaf picture. This research was uniquely detecting the leaf disease. It was based on two approaches, where circular based scheme discover the characteristics of the leaf picture. After identifying the characteristics, probable neural network method was applied to recognize the existing infection on leaves. This research was utilized to recognize the well and diseased leaf on the basis of the feature area recognition. This research analysed in random collection of the leaf pictures from sites meant for various plant leaves. Experimental analysis was done to identify the infected leaf. **Padol, P. B. and Yadav, A. A et al., 2016[21]** aimed at recognition and

arrangement of the infected leaf utilizing support vector machine method. K mean cluster approach was applied for segmenting of extracted characteristics of color and texture. In the final approach, accuracy was acquired up to 89% for detection and classification of the leaf. **Suresha, M., Shreekanth, K. N. and Thirumalesh, B. V. et al., 2017 [22]** proposed a research on technique for recognition of the blast and brown point leaf disease. Globalized and k nearest neighbours technique had been analysed for the classification of the information. The classification was done on the basis of the geometrical characteristics such as region, main alignment, and negligible alignment. In this research, a fungal infection has been determined on the basis of the accuracy level.

III. RESEARCH METHODOLOGY

The main objectives of the research work using SVM and HK-means Clustering method. To study the various diseases in the leaves of the fruit plants. To design a technique to detect the disease present in orange leaf using SVM and K-mean algorithms.To implement the HOG proposed technique. To generate the results related to healthy and infected leaves.

In this section elaborate the research methodology flow chart:

The detection of disease on orange fruit plant will be initiated by acquiring the image from the dataset. The image will be pre-processed with the different image processing techniques such as region based, edge based, model based [1]. After the image pre-processing phase, image segmentation will be performed. Segmentation is done using boundary detection and spot detection algorithm. K-Mean clustering will be used for image segmentation. HOG method has implemented to extract the global type of feature sets, calculate the number of image pixels and feature vector. Classification method has used to detect the disease in orange leaves and identify the categories. Evaluate the performance parameters such as Accuracy, Precision, Recall, FAR and FRR. Comparison.

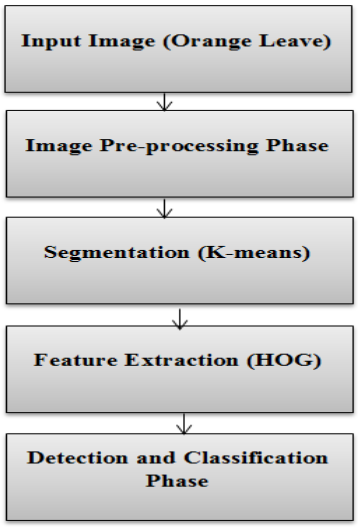
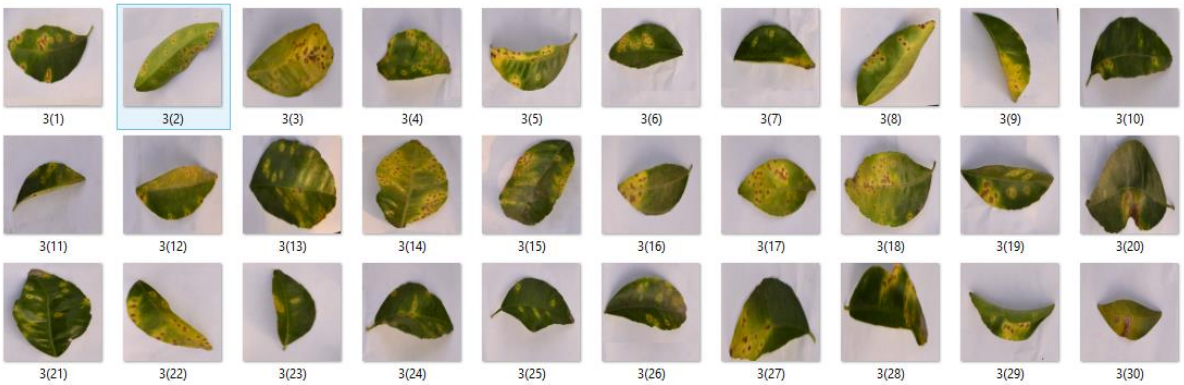


Fig.2: Proposed Flow chart

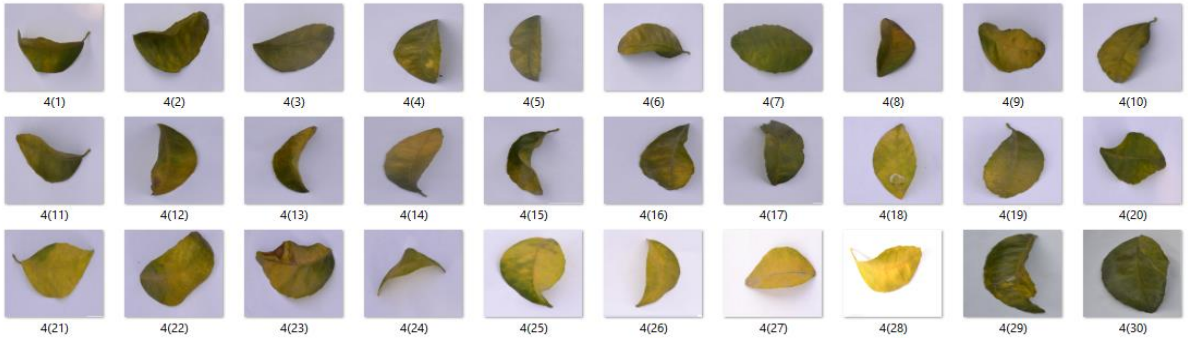
IV. EXPERIMENT RESULT

Plant disease is in charge for decreasing the fabrication that lead to economical loss. Citrus is utilised as main resource of nutrition such as vitamin C among the world. Moreover, citrus disease has an adverse effect on the manufacture and the quality of the citrus fruit. The database consist a picture album of healthy and unhealthy plant and leaves that may be practical for the scientists to avoid plant from infection utilising developed computerised visualisation method. The diseases are blackspot, canker, melanose and greening [23].



(i) Canker Database

(ii) Black Spots



(iii) Green Leaves

(iv) Normal Leaves

Fig.3: Dataset Citrus Orange leaves

The database consists of 750 pictures of fit and unhealthy pictures for both citrus plant and leaves. Every picture consists of 256* 25 sizes with 72 dpi resolution.

Training and Testing Section

In the training section is also called a knowledge domain system and Testing Section is also called analysis system. In training section, to train the multiple images one at a time, applying image pre-processing steps or methods, extract the features (HOG), segmentation technique (K-means) and classification done by SVM machine learning method. In Test section, upload the test image. Convert the original image into grayscale image. To identify the noises in the uploaded testing image, Smooth image calculated by Median Filter. Extract the edges and Features using HoG method. After that we implement testing analyser with SVM using RGF method to classify and detect the diseases in the orange leaves. Evaluate the performance Parameters based on the FAR, FRR and Accuracy Rate and compared with the existing methods.

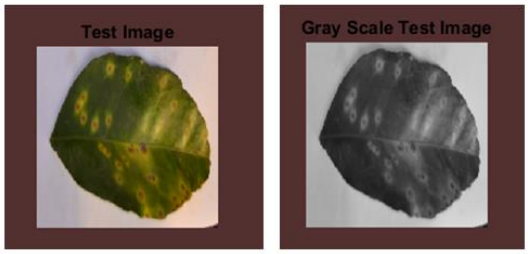


Fig.4: (i) Test Image and (ii) Gray Scale Image

Above figure 4(i) and (ii) has shown upload the test image which is a colour image. Convert the color image to grayscale image to reduce the image pixels.

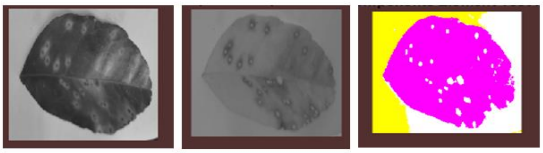


Fig.5: (i) Y component (ii) Cb component and (iii) Lab space Model

Fig 5 (i) and (ii) shown the various color space models. YCbCr color model based on detection of the color components. It is widely used for digital video. This color model format (Luminance) is saved as a single color (Red) component and chromatic information is saved as two-different components CB and CR. (i) Cb defines the difference among the blue color component and reference value. And (ii) Cr defines the difference among the red color component and a ref value. Fig 5 (iii) lab model implements to fetch the image color components red, green and blue one at a time. It is the most representative of color components not normally used. It is normally converted to minimum accuracy color spaces like as RGB (Red, Green and Blue) and CYMK is a process of four color components.

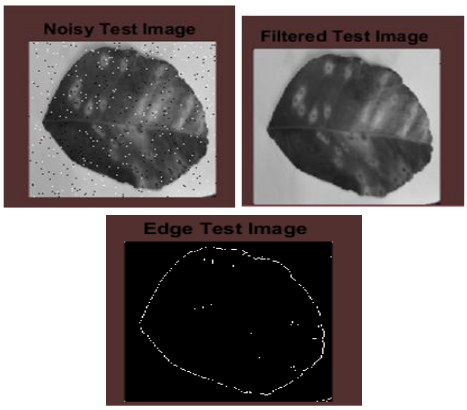


Fig.6: (i) Noisy Image and (ii) Filter Image (iii) Edge Image

Above figure 6(i) has shown noise image using Salt and Pepper Noise. It identifies the distortion of the orange leave image. 6(ii) Filtration process used to remove the noise in the noisy image using MEDIAN or 2D Transformation Method. It calculates the smooth image. 6(iii) edge detection with the Prewitt operator. It generates a picture emphasizing edges calculating an approx... of image gradient and integer-valued filter in the directions (verticals and horizontal) and is therefore relatively inexpensive in terms of calculations.

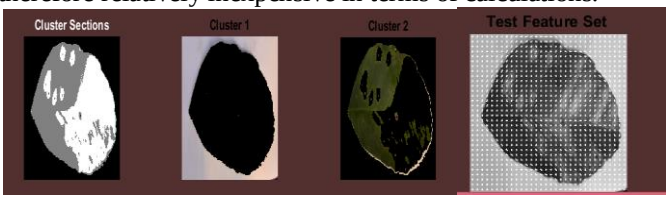


Fig7: (i) Cluster Image and Extracted Image

Fig 7 (i) K mean clustering is the method of grouping pixels into different k groups which is called as clusters and every cluster have an individual pixel. K –mean is the clustering approach used an unsupervised approach and vector quantization. Generally, the collection of the objects is the same among them and different to objects related to clusters. K-mean clustering is segmentation method based on clustering approach. Fig 7(ii) Histogram of oriented gradients is description of characteristics utilizing in computer and digital image processing for detection of object image. The method contains the arrangement of the gradients in local parts of the picture. The main feature of the histogram oriented gradients is the native look and pattern of the object determined through boundaries object. The pictures are segmented into smaller linked areas and pixels present in every group where histogram of gradient is gathered.

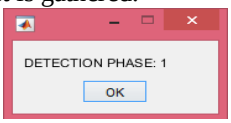


Fig.8: Detection Phase

The detection phase using Support Vector Machine Approach. In this classification approach to detect and analyze the orange leaves. Its train the complete knowledge using (i) Normal Leaves (ii) Black Spot Leaves (iii) Canker Leaves and (iv) Green Leaves.

Fig.9: Comparison – Recall

The above figure 9 demonstrates the comparison between HK-SVM and SVM classifier. Both the algorithm are compared that determines HK-SVM with maximum recall rate.

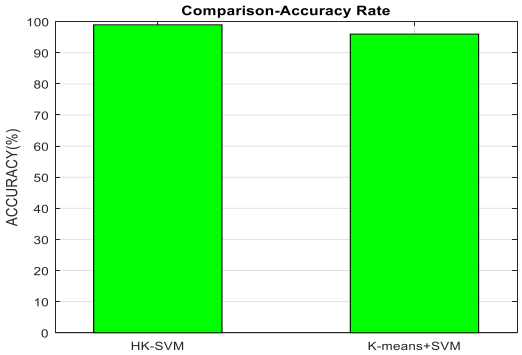


Fig.10: Comparison –Accuracy Rate

The above figure 10. demonstrates the comparison between HK-SVM and SVM. Both the algorithm are compared that determines HK-SVM with maximum accuracy rate as compared with the existing SVM algorithm.

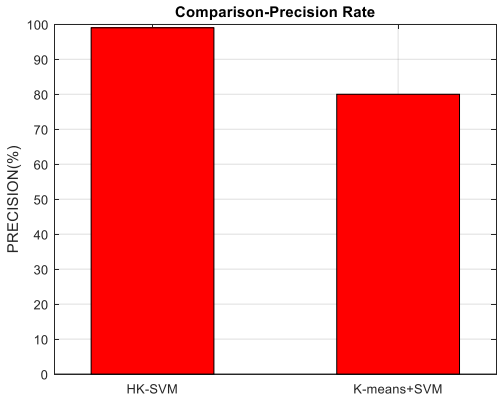


Fig.11: Comparison – Precision Rate

The above figure 11. demonstrates the comparison between HK-SVM and SVM classifier. Both the algorithm are compared that determines HK-SVM with maximum precision rate.

Table 1: Comparison between with Proposed and Existing methods

Parameters	**HK-SVM**	**SVM**
Accuracy rate	98%	96%
Recall	0.98	80
Precision	0.99	95

Table 1: shows the comparison between proposed (HK-SVM) and Existing (SVM) methods with Accuracy, Precision and Recall Value.

Table 2: Comparative Analysis with Various Test Cases.

Image	**Accuracy rate**	**Precision**	**Recall**	**Detection Orange Leave**
	97.2	0.982	0.97	Normal Leave
	98.0	0.99	0.98	Black Spot
	98.1	0.992	0.982	Canker leave
	97.3	0.983	0.976	Green Leave

Table 2: shows the comparative analysis with the various detection orange leaves based on the performance parameters such as Accuracy Rate, Precision rate and Recall. It has improved the accuracy rate and reduce the wrong leave acceptance and rejection rate (FAR and FRR).

V. CONCLUSION AND FUTURE SCOPE

The proposed technique is concluded in the MASTER's thesis improve the performance of an expert system to detect and identify the fruit leave diseases. In this approach depends on DIP to the first detection and classify the orange leaves according to the diseases is used. In this research work, collection of the various types of categories in the orange leaves such as green, healthy, black spot and canker leaves. The knowledge base design by the proposed method, image pre-processing phase has implemented to identify the error in the uploaded image, smooth image calculated, region detection to detect the background and image region in the orange leave image and convert the rgb3gray scale format. Research work has implemented a HOG method which is used to extraction of the features. HOG method to identify the unique or genuine properties of the uploaded images and obtain the features of the disease symptoms. K-means clustering method used to segment the image. In clustering method to divide the image into sub-cluster section. SVM classification has implemented to classify the disease in the

orange leaves based on the Kernal tricks. In Kernal tricks to calculate the distance of the classes. In this proposed method to resolve the detection problem in orange leaves disease. In research proposal phase, SVM is implemented with HOG and K-means clustering to enhance the disease detection results. The Graphical User Interface system for calculated consequences and accuracy are shown in the accuracy rate of the proposed combined concept changes from 96% to 98%. The comparative analysis of the image processing method is better as compared to SVM and HK-SVM method. So, the proposed image processing concept is effective to consider and calculate the plant diseases. The performance of existing method was calculated by its accuracy rate. The proposed method to improve the performance is calculated by its accuracy rate, precision, recall, false acceptance rate and false rejection rate.

In the future, the research methodology can be integrated with Ant Lion Optimization Approach to select the extracted feature with the help of fitness function. It will implement a classification and detection using color space and texture analysis to designed a system for early plant disease prediction and performance analysis of the system can enhanced in the future by using background division approaches to divide the fruit leave object from a difficult background and Improve the image Quality Factor.

VI. REFERENCES

[1]. Sankaran, S., Mishra, A., Ehsani, R., & Davis, C. (2010). A review of advanced techniques for detecting plant diseases. *Computers and Electronics in Agriculture*, *72*(1), 1-13.

[2]. Khirade, S. D., & Patil, A. B. (2015, February). Plant disease detection using image processing. In *2015 International conference on computing communication control and automation* (pp. 768-771). IEEE.

[3]. Sannakki, S. S., Rajpurohit, V. S., Nargund, V. B., & Kulkarni, P. (2013, July). Diagnosis and classification of grape leaf diseases using neural networks. In *2013 Fourth International Conference on Computing, Communications and Networking Technologies (ICCCNT)* (pp. 1-5). IEEE.

[4]. Jhuria, M., Kumar, A., & Borse, R. (2013, December). Image processing for smart farming: Detection of disease and fruit grading. In *2013 IEEE Second International Conference on Image Information Processing (ICIIP-2013)* (pp. 521-526). IEEE.

[5]. Revathi, P., & Hemalatha, M. (2012, July). Advance computing enrichment evaluation of cotton leaf spot disease detection using Image Edge detection. In *2012 Third International Conference on Computing, Communication and Networking Technologies (ICCCNT'12)* (pp. 1-5). IEEE.

[6]. Rosenzweig, C., Iglesias, A., Yang, X. B., Epstein, P. R., & Chivian, E. (2001). Climate change and extreme weather events; implications for food production, plant diseases, and pests. *Global change & human health*, *2*(2), 90-104.

[7]. Patil, S., & Chandavale, A. (2015). A survey on methods of plant disease detection. *International journal of Science and Research (IJSR), 4*, 1392-1396.

[8]. Golhani, K., Balasundram, S. K., Vadamalai, G., & Pradhan, B. (2018). A review of neural networks in plant disease detection using hyperspectral data. *Information Processing in Agriculture*, *5*(3), 354-371.

[9]. Ferentinos, K. P. (2018). Deep learning models for plant disease detection and diagnosis. *Computers and Electronics in Agriculture*, *145*, 311-318.

[10]. Martinelli, F., Scalenghe, R., Davino, S., Panno, S., Scuderi, G., Ruisi, P., ... & Davis, C. E. (2015). Advanced methods of plant disease detection. A review. *Agronomy for Sustainable Development*, *35*(1), 1-25.

[11]. López, M. M., Bertolini, E., Olmos, A., Caruso, P., Gorris, M. T., Llop, P., ... & Cambra, M. (2003). Innovative tools for detection of plant pathogenic viruses and bacteria. *International Microbiology*, *6*(4), 233-243.

[12]. Weizheng, S., Yachun, W., Zhanliang, C., & Hongda, W. (2008, December). Grading method of leaf spot disease based on image processing. In *2008 international conference on computer science and software engineering* (Vol. 6, pp. 491-494). IEEE.

[13]. Ramakrishnan, M. (2015, April). Groundnut leaf disease detection and classification by using back probagation algorithm. In *2015 International Conference on Communications and Signal Processing (ICCSP)* (pp. 0964-0968). IEEE.

[14]. Gaikwad, V. P., & Musande, V. (2017, October). Wheat disease detection using image processing. In *2017 1st International Conference on Intelligent Systems and Information Management (ICISIM)* (pp. 110-112). IEEE.

[15]. Khirade, S. D., & Patil, A. B. (2015, February). Plant disease detection using image processing. In *2015 International conference on computing communication control and automation* (pp. 768-771). IEEE.

[16]. Sunny, S and Gandhi, M. I. (2018). An efficient citrus canker detection method based on contrast limited adaptive histogram equalization enhancement. *Int. J. Appl. Eng. Res.*, *13*(1), 809-815.

[17]. Singh, V and Misra, A. K. (2017). Detection of plant leaf diseases using image segmentation and soft computing techniques. *Information processing in Agriculture*, *4*(1), 41-49.

[18]. Barbedo, J. G. A. (2013). Digital image processing techniques for detecting, quantifying and classifying plant diseases. *SpringerPlus*, *2*(1), 660.

[19]. Ashourloo, D., Aghighi, H., Matkan, A. A., Mobasheri, M. R and Rad, A. M. (2016). An investigation into machine learning regression techniques for the leaf rust disease detection using hyperspectral measurement. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *9*(9), 4344-4351

[20]. Soni, P and Chahar, R. (2016). A segmentation improved robust PNN model for disease identification in different leaf images. In *2016 IEEE 1st International Conference on Power Electronics, Intelligent Control and Energy Systems (ICPEICES)* (pp. 1-5). IEEE.

[21]. Padol, P. B and Yadav, A. A. (2016). SVM classifier based grape leaf disease detection. In *2016 Conference on advances in signal processing (CASP)* (pp. 175-179). IEEE.

[22]. Suresha, M., Shreekanth, K. N and Thirumalesh, B. V. (2017). Recognition of diseases in paddy leaves using knn classifier. In *2017 2nd International Conference for Convergence in Technology (I2CT)* (pp. 663-666). IEEE.

[23]. Sharif, M., Khan, M. A., Iqbal, Z., Azam, M. F., Lali, M. I. U. and Javed, M. Y. (2018). Detection and classification of citrus diseases in agriculture based on optimized weighted segmentation and feature selection. *Computers and electronics in agriculture*, vol *150(3)*, pp. 220-234.