JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

# AN IMPLEMENTATION OF MATLAB BASED SYSTEM FOR DISEASE AND SEVERITY DETECTION OF CROP AND PREDICTION

**Gayatri Rallabandi[1], Prof. Namrata Verma[2] , Prof. Dr. Pankaj Kumar Mishra[3]**

*[1](M-Tech Scholar), Department of Digital Electronics, Rungta College of Engineering and Technology, Kohka Road Kurud Bhilai (C.G)*
*[2]Assistant Professor, Department of Electronics and Telecommunication, Rungta College of Engineering and Technology, Kohka Road Kurud Bhilai (C.G)*
*[3]Professor, Department of Electronics and Telecommunication, Rungta College of Engineering and Technology, Kohka Road Kurud Bhilai (C.G)*

---***---

**Abstract—**Each year, plant's viral and fungal infections cause up to 30 percent of crop damage and cause spread of deadly diseases such as cancer, autism and asthma due to overuse of pesticides. Therefore, it is important to detect the disease at an early stage in the plant. Current methods involve expert eye examination. It also involves constant monitoring, which is a very expensive and difficult exercise for large farms. Farmers in India are less educated to recognise the disease and its dire consequences. They neglect or use pesticides carelessly to get out of the problem. It not only spreads deadly diseases but also contaminates the soil, also it is not economically viable for them. Therefore, it is important to recognize the diseases and their proper treatment for good crop production and environmental balance.The detection of plant leaf diseases by automated method is useful because it reduces the large effort of observation in large fields and it detects the symptoms of diseases at an early stage. Detection of plant leaf disease involves steps such as image acquisition, image pre-processing, image segmentation, feature extraction and classification. This paper implements a technology for image pre-processing. The image segmentation algorithm is used for automatic detection and research on various plant leaf disease and classification algorithms is used for leaf disease classification.

***Keywords***- KNN, Plant diseases detection.

## I. INTRODUCTION

It is estimated that there are over 1.7 million species living on earth. Of this huge amount of living algae, animals, plants and people, only plants of various species are important to create a life cycle of human life. Plants provide a way of life for humans in the form of oxygen and other important resources. Many medicines and dietary supplements are the gifts of plants to the humans. Therefore, plants are essential to human life and need to be protected at every stage.

Plants play an important role in working and maintaining balance on this earth. As humans have evolved in the technological age, we have devised many ways to feed the entire population. In the latter part of the 20th century, with the advancement of allied technology we are able to sustain the food demands of the entire human population [2] but this pressure was unbearable for the whole of civilization. These estimates put a very strong pressure on agriculture, which is still the primary source of food production for humans. Polluted and diseased plants made the situation worse.

Nowadays, photo processing methods are suitable, efficient and reliable methods for diagnosing diseases with the help of images of plant leaves. Farmers need quick and effective methods to diagnose all plant diseases. These programs can reduce loss of produce and the use of pesticides. In order to improve agricultural yields different ideas are suggested by scientists with the help of laboratory and plant diagnostic programs. The paper we have presented here researches different types of plant diseases and disease diagnostic techniques by different researchers. Image processing, threshold algorithm, K-Means cluster techniques are used for identification disease from infected plant leaves and subsequent clustering. It helps in developing various disease control strategies to help the agricultural sector [2]. Support vector machines and digital image processing have been used on leaf specimens to identify the plant as healthy or unhealthy, as well as to identify the type of disease. Such methods help the farmer to get proper treatment of diseases and give results on time. [3]

**Objective**

The disease detection model will help the users to identify a disease in a plant and will suggest the user amount and the type of pesticides a user must add to fight with the disease present in the plant.

Given a large set of data, plant diagnostics are made by comparing different plant details with those of a given plant in the data set, the disease is determined by baseline value and type of variance between data set values and actual values.

Disease detection, crop sensitivity and pesticide prediction will use different algorithms for different types of proposal regarding plant condition.

- Provide access to a very large information collection
- Provide User-friendly interface
- Online Supports Advance Features
- Suggests amount and type of fertilizers to be added in the crop.
- The Interface has to be simple to use, as the target end-users for the system
- Easy to handle by any non-technical persons.

We try to implement a Web based technology with MATLAB, where farmer can upload their plant leaf images and after software analysis doctor can easy upload their recommendations/suggestions like how much fertilizer should be used in a plant.

## II. PROBLEM IDENTIFICATION

Plant leaf disease identification is very crucial as it reduces the quality and value of the crop yield. The identification and collection of plant leaf disease is an important part of improving agricultural efficiency and financial development of producers. Acquisition and characterization is one of the most interesting and talked topics in design and IT industry.

The diagnosis of plant leaf disease is very difficult in stages. Here, the disease has affected the plant databases; the leaves are considered to suffer from alternaria alternata , anthracnose, bacterial blight & cercospora leaf spot. The main purpose of this paper is to determine whether the disease is affected in some part of the leaf or the healthy part of the leaf. We then calculate the percentage of the affected parts of the leaves by their separation. The overall results were assessed on the nature of the method.

## III. METHODOLOGY

In methodology, the basic steps for plant disease detection and classification using image processing are shown (Fig. 1).

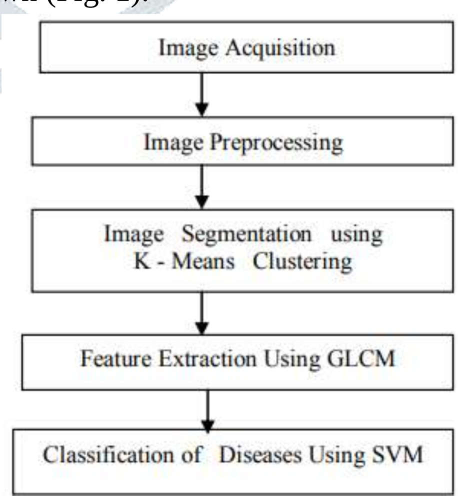


Figure 1 – Overview of the proposed Method

***Image Acquisition***

First, images of various leaves are captured using a digital camera with the required resolution. The input image has been updated to 256x256 pixels. The creation of a database depends on the application required. The image website should be carefully constructed because it usually determines the efficiency of the partition and the performance of the proposed method.

***Image Pre-Processing***

Images in the database have a different background light that affects the accuracy of the application. Pre-image processing is important for sound removal and image separation which helps to improve the accuracy of the KNN model. Distinction is made that produces only the right part of the image. Therefore, after performing the classification, all images of leaves with a dark background are available. In addition, managing isolated lighting conditions has been converted to gray images and is being transferred to further development. Pre-processed images are reduced image size and image crop to a given input. It processes and enhances the image to its needed color scale. Pre-image processing is used to improve the image quality required for further processing which includes color space adjustment and image enhancement. RGB leaf images are converted into L * a * b * color space. Color modification is done to determine the light layers and chromaticity. Color space conversion is used for the development of visual analysis.

***Image Segmentation***

Image segmentation is the process used to simplify the representation of an image into meaningful form, such as to highlight objects of interest from background. The K-means clustering algorithm performs segmentation by minimizing the sum of squares of distances between the image intensities and the cluster centroids. K-means clustering algorithm, or Lloyd's algorithm, is an iterative algorithm that partitions the data and assigns observations to precisely one of k clusters defined by centroids.

The steps in the algorithm are given below.

1) Choose k initial cluster centers (centroid).
2) Compute point-to-cluster-centroid distances of all observations to each centroid.
3) Assign each observation to the cluster with the closest centroid.
4) Compute the mean of the observations in each cluster to obtain k new centroid locations.
5) Repeat steps 2 through 4 until there is no change in the cluster assignments or the maximum number of iterations is reached.

***Feature Extraction***

After splitting, the GLCM features are extracted from the image. The Gray-Level Co-Occurrence Matrix (GLCM) is a mathematical method of investigating texture that looks at the pixel-local relationships [15]. GLCM functions define image formation by incorporating spatial relationships between pixels in images. Mathematical steps are extracted from this matrix. In creating GLCMs, a list of descriptions that define the directional pixel relationship with a different distance will be specified. In this proposed approach, four elements are extracted which include brightness, strength, similarity and integration. Let Pij represent input (i, j) in the standard Gray-Level CoOccurrence Matrix. N represents the number of different gray levels in the measured image.

***SVM (Support Vector Machine)***

Support Vector Machine is a kernel-based supervised learning algorithm used as a classification tool. The SVM training algorithm increases the gap between training data and class boundaries. Paid decision work depends only on training data called as support veggies, which are very close to the decision boundary. It works best in a large size area where the size is greater than the training data number. SVM converts data from input space to high-level space using kernel function. Nonlinear data can also be separated using a hyperplane at a higher size. Computer complexity is reduced by the kernel Hilbert space (RKHS).

Basically a straight forward model of reversal and segregation problems. It can find a solution to both direct and indirect problems and works well with many functional problems. The SVM method is like this: the algorithm develops a line or hyper plane that divides the data into many different categories. On the first scale what you do is find a line (or hyper plane) that separates the data of those two categories.

The supporting vectors of those data points are very close to the hyperplane and influence the position and position of this hyperplane. Using supporting vectors, we try to enlarge the dividing line. When you remove these support vectors, the position of the hyper plane will change.

### *Classification Of The Disease*

The classification of the disease is done in two steps. The first step is to determine the type of crop and the second step is to determine the type of disease. To perform these tasks using convolutional neural networks. Learning transfer is used to build a deeper learning model and is trained using the ImageNet database. Learning transfer is a form of machine learning in which a model is trained in one task and transferred to another related task. It is a method by which pre-trained neural networks are used to construct neural network with the same type of work to incorporate faster and more stable problem-solving progress. These imaginative networks are built with training on large databases that contain a very large number of different images. Several research organizations are developing such type of models that take weeks to train with the latest high-end hardware. This is issued under a permitted license for direct re-use to create a new problem-solving model. These previously trained types can be optimized using a new database, if its nature is similar to the database in which the network is trained. In such cases the training is only trained in the last layer of the network, the set network can be used directly to solve the problem. If the database size is large enough the previously trained model can be re-trained using the new data again, in that case, the neural network is initiated by the weights of the selected model.

## IV. Test and Implementation

This modern system is connected with an online web platform. It contains mainly four sections : farmer data uploading section, leaf data download technical operator section, disease analysis section and last agrologist section.

### A. *Farmer Registration and Technical Operator Section*

In this section figure 2(a) farmers can upload plant leaf photos using a mobile browser, and register their name. This is easy to access UI interface for farmers.


**(a) (b)**

Figure 2-(a) Farmer Registration (b) Technical Operator

In this figure 2(b), a technical operator which is available in the lab, they download the data from respected farmer images. They store data in the folder of the testing area in software.

### B. *Plant Disease Analysis Section*

In this processing section mainly three works are done: image acquisition , pre-processing and classification.

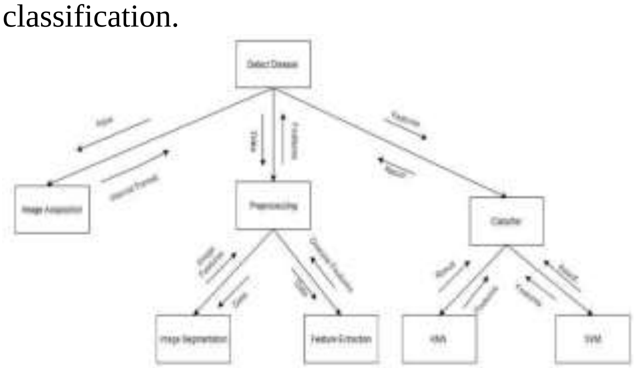

Figure 3- Software Distribution System design

### *i. Image Acquisition*

It is the initial step in which image is uploaded in the system for detecting the Disease segment in the image. The image can be uploaded in any format like jpeg, png ,etc.

### ii. Image Pre-processing

- ***Image Segmentation***

  The image segmentation means to segment each pixel of the image. A matrix will be formed in which the value of each pixel of image will be stored in the form of RGB (Red, Green, Blue) value. All pixels will be assigned a value by the image segmentation.

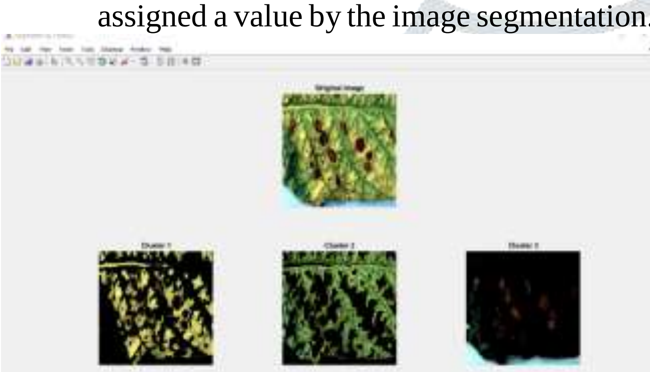

Figure 4-Cluster 1,2 & 3 are Segmented

- ***Image Features Extraction***

  The output feature is used to reduce the original data sets and to create new subsets of features with a combination of existing features. In the implemented work, a text element, a shape feature and an effective pixel value are extracted. Texture, texture and pixel value are extracted using GLCM. Feature identification plays an important role in identifying objects. Feature extraction is used in many image processing applications. Some of the most important factors used to diagnose plant diseases are color, texture, morphological, edge etc. among all these factors the most active factor is morphological. It gives a very good result. In this process we use the color combination method.

  Gray Level Co-occurrence Matrix (GLCM)
  The Gray Level Co-occurrence Matrix is used for the extraction of the texture element especially in image processing [15]. The area of each gray pixel can be displayed in GLCM methods. In this work, a total of 11 synthetic materials were calculated using GLCM methods [16]. Features released at GLCM are Mean Standard Deviation, Energy, Entropy, Skewness, Variance, Kurtosis, Smoothness, Contract and Correlation. In the following table I show the features of GLCM and formulas.

| Features | Formulas |
|---|---|
| Mean | $\mu = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N} p(i,j)$ |
| Standard Deviation | $\sigma = \sqrt{\frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}(p(i,j)-\mu)^2}$ |
| Energy | $E = \sum_{i,j} p(i,j)^2$ |
| Entropy | $h = -\sum_{k=0}^{l-1} prk(log_2 prk)$ |
| Skewness | $S = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}\left(\frac{(p(i,j)-\mu^{z}}{\sigma}\right)$ |
| Variance | $Var = (S.D)^2$ |
| Homogeneity | $H = \sum_{i,j}\frac{p(i,j)}{1+\lvert i-j\rvert}$ |
| Kurtosis | $K = \left\{\frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}\left(\frac{(p(i,j)-\mu^{z}}{\sigma}\right)\right\} - 3$ |
| Smoothness | $R = 1 - \frac{1}{1+\sigma^2}$ |
| Contrast | $C = \sum_{i=1}^{M}\sum_{j=1}^{N}(i,j)^2 p(i,j)$ |
| Correlation | $cor = \sum_{i=1}^{M}\sum_{j=1}^{N}\frac{(i-\mu)(j-\mu)p(i,j)}{\sigma_1\sigma_2}$ |

Table 1- GLCM Features and Formulas

### *iii. Image Classification*

In this we will use three classifiers for classification of images. The classifiers are KNN, SVM,Bayesian classifier. These classifiers are used to classify the image. We will get output foreach classifier and we will compare the result of each classifier and decide which will give the most optimum result.

- ***SVM Classifier***

  The following function takes training data,training labels and test data to assign labels to test data. The code determines the type of disease using the Multi-Class Support Vector Machine algorithm. Multi-class classification is an important and ongoing research subject in machine learning and data mining.For k-class problem, this method constructs k classifiers, where each one is trained on data from one class.

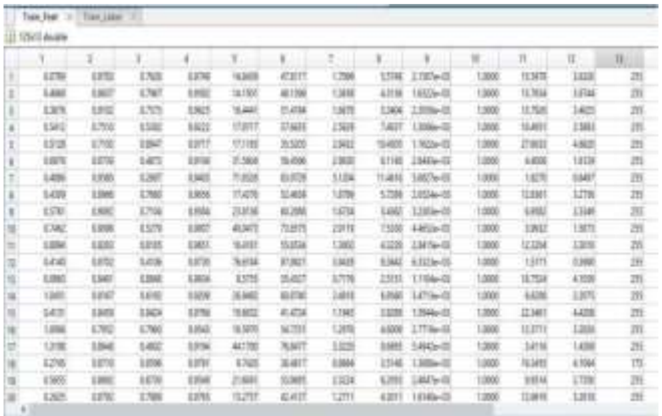

Figure 5 – Training Data .Mat file

- ***KNN Classifier***

The following function takes nearest neighbours, training data,training labels and test data to assign labels to test data. The code determines the type of disease using the K-nearest neighbours algorithm. K nearest neighbors is a simple algorithm that stores all available cases and classifies new cases based on a similarity measure (e.g., distance functions). In this case the distance function used is Euclidean Distance. KNN has been used in statistical estimation and pattern recognition already in the beginning of 1970's as a non-parametric technique.

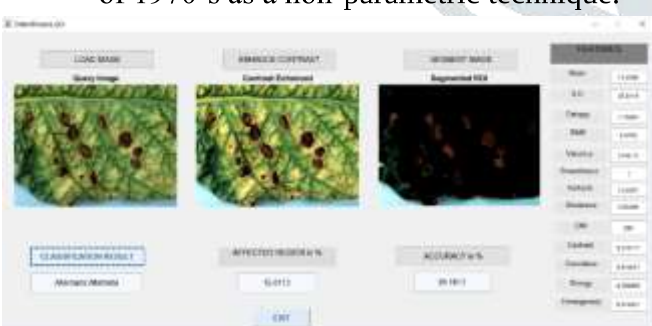

Figure 6- Gui Model of our final system

A case is classified by a majority vote of its neighbors, with the case being assigned to the class most common amongst its K nearest neighbors measured by a distance function. If K = 1, then the case is simply assigned to the class of its nearest neighbor.

### *C. Doctor/ Agrilogist Result, Precaution Submission Section*

In section Matlab UI specialist can upload the details of precaution , and how to control via fertilizer and other ways. It's connected to the web platform where farmer can be notify the details.


Figure 7- Specialist Result Upload Section

## V. Result and Discussion

The method of focusing on image processing is applied to the automatic leaf disease classification established on leaf image processing. Since the images are captured directly from the field by the farmers without much effort, the project system can be applied using practical requirements. The system approach advises the farmer with minimal effort. It requires the most effective farmer to capture the leaf image of the plant using a mobile camera and forward it to the DSS without additional input.

In the implemented system, a database consisting of 90 images of grape leaves is created manually. Training dataset includes 30 images for healthy leaves and 30 images for unhealthy leaves. Testing dataset includes 30 images including healthy and unhealthy leaves. Matlab tool is used to train the images using 22 combined features (4 texture and 18 color features). An image is randomly selected from the database for testing.

## VI. Conclusion and Future Work

To achieve the type of final disease, we have used k-means clustering for the segmentation of the image, GLCM for texture extraction, means mean for the release of a color feature with KNN, Neural Network, SVM for the final separation of the leaf disease. The above research on the diagnosis of plant diseases using the process of image processing is carried out using important tools such as the integration of k-mean, GLCM. Some of the challenges we have to address plant diseases very quickly. In the first stage of the disease we must get the disease, for this we must monitor the plant continuously. In future work we use a large amount of data as the purpose of training in the neural network. As we increase the training data the accuracy of the system will be higher. And in the future we can use abstract concepts as tools of division and we can compare the accuracy and speed of the system. K-Means algorithm provides very high accuracy, but is useful for one database at a time. While the neural network is useful for more information, once it has been trained. Neural network also provides good accuracy.

In the future different neural network algorithms may be used to classify diseased leaf images and the segmentation effects of those images will be matched.

with the results of existing methods of separation. In the future the work may be expanded by increasing the number of images and the type of leaf diseases. Also, we can use a variety of classification, genetic engineering techniques to detect and classify leaf diseases.

## VII. Acknowledgment

Expression of giving thanks is just a part of those feelings which are too large for words but shall remain as memories of beautiful people with whom I have got the pleasure of working during the completion of this work. I am grateful to Prof. Namrata Verma, Bhilai, (C.G.), who helped me to complete my work by giving an encouraging environment. I want to express my truthful gratitude to Prof. Dr. Pankaj Kumar Mishra. His comprehensive knowledge and his logical way of thinking have been of great value to me. His understanding, encouragement, and personal guidance have provided a sound basis for the present work.

## 6) References

[1] M. H. Sheikh, M. S. Reza and K. Fatema, "Plant Leaf Diseases Detection Using Image Processing System," *2019*.

[2] K. Gayatri, R. D. Kanti, V. C. S. R. Rayavarapu, B. Sridhar and V. R. G. Bobbili, "Image processing and Pattern Recognition based Plant Leaf diseases Identification and Classification," *Journal of Physics: Conference Series*, vol. 1804, no. 1, p. 12160, 2021.

[3] N. Nisar, A. Awasthi, M. Chhabra and A. I. Abidi, "Image based Recognition of Plant Leaf Diseases: A Review," in *2020 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)*, 2020.

[4] V. Pooja, R. Das and V. Kanchana, "Identification of plant leaf diseases using image processing techniques," in *2017 IEEE Technological Innovations in ICT for Agriculture and Rural Development (TIAR)*, 2017.

[5] Y. H. S. Kumar and G. Suhas, "Identification and Classification of Fruit Diseases," *International Conference on Recent Trends in Image Processing and Pattern Recognition*, pp. 382-390, 2016.

[6] K. S. Devi, P. Srinivasan and S. Bandhopadhyay, "H2K – A robust and optimum approach for detection and classification of groundnut leaf diseases," *Computers and Electronics in Agriculture*, vol. 178, p. 105749, 2020.

[7] F. Qin and H. Wang, "Development of a Recognition System for Alfalfa Leaf Diseases Based on Image Processing Technology," in *10th International Conference on Computer and Computing Technologies in Agriculture (CCTA)*, 2016.

[8] V. Singh and A. Misra, "Detection of plant leaf diseases using image segmentation and soft computing techniques," *Information Processing in Agriculture*, vol. 4, no. 1, pp. 41-49, 2017.

[9] A. S. Rajput, S. Shukla and S. S. Thakur, "Detection Of Plant Leaf Diseases In Agriculture Using Recent Image Processing Techniques – A Review," *International Journal of Students' Research in Technology & Management*, vol. 7, no. 1, pp. 29-33, 2019.

[10] N. Kaur, V. Devendran and S. Verma, "Detection of Plant Leaf Diseases by Applying Image Processing Schemes," *Journal of Computational and Theoretical Nanoscience*, vol. 16, no. 9, pp. 3728-3734, 2019.

[11] R. M. Prakash, G. Saraswathy, G. Ramalakshmi, K. Mangaleswari and T. Kaviya, "Detection of leaf diseases and classification using digital image processing," in *2017 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS)*, 2017.

[12] N. Ananthi, K. Kumaran, M. V, G. M. S and H. P, "Detection and Identification of Potato Plant Leaf Diseases using Convolution Neural Networks," *European Journal of Molecular & Clinical Medicine*, vol. 7, no. 4, pp. 2753-2762, 2020.

[13] N. B. Wadekar, P. K. Sharma and N. S. Sapkale, "Detection and Controlling of Grape Leaf Diseases using Image Processing and Embedded System," *International Journal of Electrical and Electronics Engineering*, vol. 2, no. 10, pp. 13-15, 2015.

[14] S. Ananthi and S. V. Varthini, "Detection And Classification Of Plant Leaf Diseases," *2012*.

[15] D. A. Bashish, M. Braik and S. Bani-Ahmad, "Detection and Classification of Leaf Diseases using K-means-based Segmentation and Neural-networks-based Classification," *Information Technology Journal*, vol. 10, no. 2, pp. 267-275, 2011.

[16] I. Kaur, G. Aggarwal and A. Verma, "Detection and Classification of Disease Affected Region of Plant Leaves using Image Processing Technique," *Indian journal of science and technology,* vol. 9, no. 48, 2016.

[17] S. R. Dubey and A. S. Jalal, "Detection and Classification of Apple Fruit Diseases Using Complete Local Binary Patterns," in *2012 Third International Conference on Computer and Communication Technology*, 2012.

[18] A. Agarwal, A. Sarkar and A. K. Dubey, "Computer Vision-Based Fruit Disease Detection and Classification," , pp. 105-115, 2019.

[19] D. Parmar, . Kathiriya and K. Patel, "Classification of citrus leaf diseases by image processing based on texture (Statistical) related features," *International Journal of Chemical Studies,* vol. 6, no. 5, pp. 3053-3061, 2018.

[20] M. S.Abirami, "Application of Image Processing in Diagnosing Guava Leaf Diseases," *International Journal of scientific research and management,* vol. 5, no. 7, pp. 5927-5933, 2017.

[21] G. Tripathi and J. Save, "An Image Processing And Neural Network Based Approach For Detection And Classification Of Plant Leaf Diseases," , 2015.

[22] R. Anand, S. Veni and J. Aravinth, "An application of image processing techniques for detection of diseases on brinjal leaves using k-means clustering method," in *2016 International Conference on Recent Trends in Information Technology (ICRTIT)*, 2016.

[23] B. S. Prajapati, V. K. Dabhi and H. B. Prajapati, "A survey on detection and classification of cotton leaf diseases," in *2016 International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT)*, 2016.

[24] H. Permuter, J. Francos and I. Jermyn, "A study of Gaussian mixture models of color and texture features for image classification and segmentation," *Pattern Recognition,* vol. 39, no. 4, pp. 695-706, 2006.

[25] D. A. Bashish, M. Braik and S. Bani-Ahmad, "A framework for detection and classification of plant leaf and stem diseases," in *2010 International Conference on Signal and Image Processing*, 2010.

[26] D. K. Parmar, "A Comprehensive Study And Analysis Of Existing Image Processing Applications In The Field Of Agriculture For Detection Classification Of Citrus Familys Plant Leaf Diseases," *University,* 2019.